

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:11:28 ; Search time 9.08923 Seconds
(without alignments)
1337.138 Million cell updates/sec

Title: us-09-685-010-47

Perfect score: 3598

Sequence: 1 MSPPKAPLKRFPNDPSCAPS.....LKEGNTNCRAPMECQESWK 725

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 106657 seqs, 16763532 residues

Total number of hits satisfying chosen parameters: 106657

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/PCTUS_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCTUS_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	374	10.4	1958	12	US-10-028-946-4
2	374	10.4	2054	12	US-10-028-946-2
3	371	10.3	2055	9	US-10-017-216-4
4	367	10.2	2053	9	US-10-017-216-2
5	358.5	10.0	2125	10	US-09-919-172-29
6	355	9.9	1597	9	US-10-017-216-6
7	354.5	9.9	868	9	US-09-884-001-19
8	353.5	9.8	677	10	US-09-745-763-168
9	350	9.7	1641	9	US-10-017-216-5
10	323	9.0	660	10	US-09-864-761-47959
11	320	8.9	1045	10	US-09-815-242-10617
12	317	8.8	2139	10	US-09-727-384-6
13	315.5	8.8	879	9	US-10-108-605-217
14	314	8.7	1179	10	US-09-815-242-13262
15	309	8.6	1179	10	US-09-815-242-13608
16	303	8.4	2310	9	US-09-991-496-120
17	303	8.4	2310	10	US-09-874-923-120
18	303	8.4	2568	10	US-09-866-108-3
19	297.5	8.3	909	10	US-09-925-299-988

20	297.5	8.3	1786	9	US-09-742-096-3	Sequence 3, Appli
21	293.5	8.2	981	10	US-09-815-242-12211	Sequence 12211, A
22	289	8.0	689	9	US-10-108-605-305	Sequence 305, App
23	287.5	8.0	1286	9	US-10-017-216-7	Sequence 7, Appli
24	286.5	8.0	751	10	US-09-864-761-38419	Sequence 38419, A
25	286	7.9	1192	10	US-09-815-242-10903	Sequence 10903, A
26	284.5	7.9	1711	10	US-09-771-161A-219	Sequence 219, App
27	284.5	7.9	1711	10	US-09-771-161A-220	Sequence 220, App
28	282	7.8	1184	10	US-09-815-242-5229	Sequence 5229, Ap
29	282	7.8	1188	10	US-09-815-242-12125	Sequence 12125, A
30	279.5	7.8	600	10	US-09-975-901-2	Sequence 2, Appli
31	279	7.8	945	10	US-09-745-763-191	Sequence 191, App
32	279	7.8	1242	10	US-09-925-299-911	Sequence 911, App
33	275.5	7.7	996	10	US-09-815-242-5251	Sequence 5251, Ap
34	275.5	7.7	1009	10	US-09-815-242-12141	Sequence 12141, A
35	272.5	7.6	1013	12	US-10-007-805-553	Sequence 553, App
36	266.5	7.4	1002	10	US-09-804-287A-475	Sequence 475, App
37	266.5	7.4	1002	12	US-10-007-805-475	Sequence 475, App
38	265.5	7.4	1095	12	US-10-007-805-493	Sequence 493, App
39	265.5	7.4	1341	12	US-10-007-805-565	Sequence 565, App
40	265.5	7.4	1349	12	US-10-007-805-573	Sequence 573, App
41	262	7.3	2835	10	US-09-885-535-4	Sequence 4, Appli
42	260	7.2	1239	12	US-10-007-805-577	Sequence 577, App
43	258	7.2	830	9	US-10-033-245-7	Sequence 7, Appli
44	258	7.2	830	9	US-10-033-223-7	Sequence 7, Appli
45	258	7.2	830	9	US-10-033-167-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1

US-10-028-946-4
; Sequence 4, Application US/10028946
; Patent No. US20020123622A1
; GENERAL INFORMATION:
; APPLICANT: Yu, Xuanchuan
; APPLICANT: Miranda, Maricar
; APPLICANT: Friddle, Carl Johan
; TITLE OF INVENTION: No. US20020123622A1el Human Kinases and Polynucleotides Encodi
; FILE REFERENCE: LEX-0289-USA
; CURRENT APPLICATION NUMBER: US/10/028,946
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/258,335
; PRIOR FILING DATE: 2000-12-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1958
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-028-946-4

Query Match	10.4%	Score 374;	DB 12;	Length 1958;
Best Local Similarity	21.2%	Pred. No. 2.9e-09;		
Matches	154;	Conservative 155;	Mismatches 297;	Indels 122; Gaps 20;
Qy	11	FNDP-----SCAPSPGAYDVKLTLEVLKGPVSFQKQKESQKQNLNVKDTTL 62		
Db	388	FDEPEKNSWSSPQLSPSGFSGEELPFV--GFSYSKALGILGRSES-----VVSGLDS 440		
Qy	63	PASARKVKSSKESQKNDKDLKLEKEIRVLLQERGAQDRRIQDLELEKMEARLNA 122		
Db	441	PA-----KTSMEKKLLIKSELQSDQKCHKMEQEMTLHRRVSEVAVLSQKEVELKA 495		
Qy	123	ALREKTSUANNATLEKOLIELTRTNELLKSKFSPNGQKNLRILSLMLKLRNRETQM 182		
Db	496	SETQSRSLLEQDLATVITECSSLRLEQARMEVSDDK-----ALQLLDIREQSRKL 549		
Qy	183	RGMAKQEGMEMKLOVTOQRSLEESQKTAQLEGLKVS-----IEKEKIDEXSET 231		
Db	550	Q-----EIKEOEQAOQVEEMRLMNQLEEDLVSAARRRSDLYSELSRESRLAAEEFK 600		

QY 232 EKLLYIEISCASDOVEYKLDIAQLEENLKEKNDIEILSKQSLEENIVILSKOVEDLN 291
DB 601 KRATECQHLLKAKDQKPEVGEYAKLEKINAEOQLKIQELOEKLE-----KAVKAST 653
QY 292 VKCQLE--KEDHVNRRNHNENLNAEMONLKQFLEQOEHEKLOOKELOIDSLQQ 349
DB 654 EATELLONRQAKERAERELEKLNREDSSEGIKKLVAEERHRSLENKVRLETMERR 713
QY 350 EKELSSSLHQKLCSPQOEEMVKEKNLFEELKOTLDLOKLOOKEQAERLVKQLEEEAKS 409
DB 714 ENRLKDDIQT-----SQIQOMADKILLEEKHREAQVSAQHLEVLHKQ 758
QY 410 R-----AELKLEELKKEALE-----KSSAAHTQATLLQEK-----YDSMVOS 452
DB 759 KQHYEEKIKVLNDQIKKDLADKETLNNMQRHEEAEHKGKILSEQKAMINAMDSKIRS 818
QY 453 LEDVTAOFESYKALTASE-----IEDLKLENSLOEKAAGNAEDVQHO 498
DB 819 LEQRIVELSEANKLAANSLFTQRMNKAQOEEMISELROOKFYLETQAGKLEAQNKLBEQ 878
QY 499 ILATESSNQEYVRMLLDLQTSALKETEITVTFLOKITDLOLQKQOEDF-----552
DB 879 LEKISHQDSDKNRLLLELETRLSVLEHEEQKLEKRLQTELQLSQOERESQLTALQNA 938
QY 553 RKQLEDE--EGRAEKENTTAELTEEINKWRLLYEELYNKTFQLOLDAFEVEKQALLNE 611
DB 939 RAALESQRLQAQTELEETAAEEETQALTAHRDEI-----QRKDALRNSCTVITD- 990
QY 612 HGAAEQOLNKIRDSYAKLLGHON--LKQIKHVVKLKDENSOLKSEVSKRLCOLAKKKOS 669
DB 991 ----LEEQLNQLTQEDNAE--LNNQNFYLSKOLDASGANDEIVOLRSEVDHLRREITEREQ 1046
QY 670 ETKLQEEEL 677
DB 1047 LTSQKQTM 1054
RESULT 2
US-10-028-946-2
; Sequence 2, Application US/10028946
; Patent No. US20020123622A1
; GENERAL INFORMATION:
; APPLICANT: Yu, Xuanchuan
; APPLICANT: Miranda, Maricar
; APPLICANT: Friddle, Carl Johan
; TITLE OF INVENTION: No. US20020123622A1el Human Kinases and Polynucleotides Encoding
; FILE REFERENCE: LEX-0289-USA
; CURRENT APPLICATION NUMBER: US/10/028,946
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/258,335
; PRIOR FILING DATE: 2000-12-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2054
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-028-946-2
Query Match 10.4%; Score 374; DB 12; Length 2054;
Best Local Similarity 21.2%; Pred. No. 3e-09;
Matches 154; Conservative 155; Mismatches 297; Indels 122; Gaps 20;
QY 11 FNDP-----SGCAPSPGAYDVKTLEVLKGPVSQKRSQRFQKQESKQNLNVDKDTTL 62
DB 388 FDEPEKNWVSSPQLSPSGFSGEELPFV--GFSYKALGILGRSES-----VVSGLDS 440
QY 63 PASARKVSSSKSKSQKNDKDLKILEKEIRVLLQERGAQDRRIQDLETELEKMEARLNA 122
DB 441 PA-----KTSSMEKKLLIKSELQDSQDKCHKMEQEMRLHRRVSEVEAVLSQKEVELKA 495
QY 123 ALREKTSLSANNATLEKOLIELTRTNELLKSKFSENGKQNLRLISLEMLKRLNKRRETM 182

DB 496 SETQSRLLLEQDLATYITECSSLSKRSLEQARMEVSEDOK-----ALQLLHDIRSQSRKL 549
QY 183 RGMMAKQEGMEMKLVQTRSLEESQGIQAQLEGKLV-----IEKEKIDEKSET 231
DB 550 Q-----EIKEQEYQAQVEEMLNMNQLDEEDLVSAARRSDLYESELRESLAAEEFK 600
QY 232 EKLLYIEISCASDOVEYKLDIAQLEENLKEKNDIEILSKQSLEENIVILSKOVEDLN 291
DB 601 KRATECQHLLKAKDQKPEVGEYAKLEKINAEOQLKIQELOEKLE-----KAVKAST 653
QY 292 VKCQLE--KEDHVNRRNHNENLNAEMONLKQFLEQOEHEKLOOKELOIDSLQQ 349
DB 654 EATELLONRQAKERAERELEKLNREDSSEGIKKLVAEERHRSLENKVRLETMERR 713
QY 350 EKELSSSLHQKLCSPQOEEMVKEKNLFEELKOTLDLOKLOOKEQAERLVKQLEEEAKS 409
DB 714 ENRLKDDIQT-----SQIQOMADKILLEEKHREAQVSAQHLEVLHKQ 758
QY 410 R-----AELKLEELKKEALE-----KSSAAHTQATLLQEK-----YDSMVOS 452
DB 759 KQHYEEKIKVLNDQIKKDLADKETLNNMQRHEEAEHKGKILSEQKAMINAMDSKIRS 818
QY 453 LEDVTAOFESYKALTASE-----IEDLKLENSLOEKAAGNAEDVQHO 498
DB 819 LEQRIVELSEANKLAANSLFTQRMNKAQOEEMISELROOKFYLETQAGKLEAQNKLBEQ 878
QY 499 ILATESSNQEYVRMLLDLQTSALKETEITVTFLOKITDLOLQKQOEDF-----552
DB 879 LEKISHQDSDKNRLLLELETRLSVLEHEEQKLEKRLQTELQLSQOERESQLTALQNA 938
QY 553 RKQLEDE--EGRAEKENTTAELTEEINKWRLLYEELYNKTFQLOLDAFEVEKQALLNE 611
DB 939 RAALESQRLQAQTELEETAAEEETQALTAHRDEI-----QRKDALRNSCTVITD- 990
QY 612 HGAAEQOLNKIRDSYAKLLGHON--LKQIKHVVKLKDENSOLKSEVSKRLCOLAKKKOS 669
DB 991 ----LEEQLNQLTQEDNAE--LNNQNFYLSKOLDASGANDEIVOLRSEVDHLRREITEREQ 1046
QY 670 ETKLQEEEL 677
DB 1047 LTSQKQTM 1054
RESULT 3
US-10-017-216-4
; Sequence 4, Application US/10017216
; Patent No. US20020160483A1
; GENERAL INFORMATION:
; APPLICANT: KAPPELLER-LIBERMANN, Rosana
; TITLE OF INVENTION: 13245, A No. US20020160483A1el Human Myotonic Dystrophy Type P
; TITLE OF INVENTION: Kinase and Uses Therefor
; FILE REFERENCE: 10147-57U1
; CURRENT APPLICATION NUMBER: US/10/017,216
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/242,429
; PRIOR FILING DATE: 2000-10-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2055
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-017-216-4
Query Match 10.3%; Score 371; DB 9; Length 2055;
Best Local Similarity 21.7%; Pred. No. 4e-09;
Matches 159; Conservative 151; Mismatches 291; Indels 132; Gaps 23;
QY 11 FNDPSG-----CAPS--PGAYDVKTLEVLKGPVSQKRSQRFQKQESKQNLNVDKDTTL 62
DB 387 FDEPEKNWVILCVPAEPFLAFSGEELPFV--GFSYKALGVLGRSES-----VVSGLDS 439
QY 63 PASARKVSSSK-----KESQKNDKDLKILEKEIRVLLQERGAQDRRIQDLETELEKME 117

```

Db 440 PA---KVSSMEKLLIKSKELQSDQKCHKMEQDMRL-----HRRVSEVEAVLSQKE 489
QY 118 ARNLAALREKTSLSANNATLEKQLIETRTNELLKSFSENGQKNRILSLMLKRNK 177
Db 490 VELKASETQSLLEQDIATYITECSSLKRSLEQARMEVSEDDK-----ALQLLHDIRE 543
QY 178 RETKMRGMAKQEGEMKQVQTSRLESQKIAQLEKGLVS-----IEKKID 226
Db 544 QSRKIQ-----EIKEQYQAOVEEMRLMNMNQLDEEDIVSARRSDLYESELRESRIA 594
QY 227 EKSTETKLYIEBISQASQVQYKLDIAOLENLEKNDIEILSKOSLEENIVILSKQ 286
Db 595 AEEFKRANECQHKLMKADQKPEVGEYSKLEKINAEOQLKIQLOEKLE-----KA 647
QY 287 VEDLVNKCALLE--KEDHVNRRNHNENLNAEMNLKQKFILOEQEHEKLOQKELQID 344
Db 648 VKASTEATELLQNRQAKERAERELEKLNHRNEDSSGIRKKLVAEERRHSLENKVRKLE 707
QY 345 SLLOQEKELSSSHQKLCSPQEEVMVKEKNLFEELKQTLDELQKLOQKEQAERLVKOLE 404
Db 708 TMRERNLKDDIOTK-----SQIQOMADKILELEEKHREAQVSAQHLE 752
QY 405 EAKSR-----AEELKLEELKKEAELEK-----SSAAHTQATLLLOEK-----YD 447
Db 753 VHLKQEKQHYEEKIKVLDNQIKKDLADKESLNMMQRHEEEAHEKGKILSEQKAMINAMD 812
QY 448 SMVQSLEDVTAQESYKALTASE-----IEDKLNSLSLOEKAAGNAE 493
Db 813 SKIRSLQEVISEANKLAANSLSLFTQRMNAQOEEMISELROQKFYLETQAGKLEAQRN 872
QY 494 DVHQIILATESSNOEYVRMLDLQTKSALKETEIKETVSLQKITDLOLQKQOEDF- 552
Db 873 KLEBQLEKISHODSHKSRILLELETRLEVSLEHEEQKLEKRLQTLQSLQRESQLT 932
QY 553 -----RKOLEDE--EGRKAQKENTAELETEINKWRLLYEELYNKTKPFLQLODAFEVEKQ 606
Db 933 ALQAAARALESQLROAKTELETTAEAEETQALTAHRDEI-----QRKFDALRNSCT 985
QY 607 ALLNEHGAEOELNKIRDSYAKLLGHON--LKOKIKHVVKLKDENSOLKSEVSKLRQOLA 664
Db 986 VITD-----LEEQLNQLTEDNAE--LNNQNFYLSKOLDEASGANDEIVOLRSEVDHLRREIT 1040
QY 665 KKKQSETKLOEEL 677
Db 1041 EREMQLTSQKQTM 1053

```

```

RESULT 4
US-10-017-216-2
; Sequence 2, Application US/10017216
; Patent No. US20020160483A1
; GENERAL INFORMATION:
; APPLICANT: KAPELLER-LIBERMANN, Rosana
; TITLE OF INVENTION: 13245, A No. US20020160483A1el Human Myotonic Dystrophy Type Pro
; FILE REFERENCE: 10147-57U1
; CURRENT APPLICATION NUMBER: US/10/017,216
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/242,429
; PRIOR FILING DATE: 2000-10-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2053
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-017-216-2

```

Query Match 10.2%; Score 367; DB 9; Length 2053;
 Best Local Similarity 21.0%; Pred. No. 6e-09;
 Matches 153; Conservative 151; Mismatches 287; Indels 136; Gaps 22;

```

QY 11 FNDP-----SGCAPSPGAYDVKTLEVLKGPVSFQSKQRFKQOKESQKQNLNVDKDDTL 62
Db 388 FDEPKNSWSSVSPQLSPSCFSGEELPFV--GFYSKALGILGRSES-----VVSGLDS 440
QY 63 PASARKVKSSEKSKNDKDLKILEKEIRVLQERGAQRRIODLLETELEKMEARLNA 122
Db 441 PA-----KTSMEKLLIKSKELQSDQKCHKMEQDMRLHRRVSEVEAVLSQKEVELKA 495
QY 123 ALREKTSLSANNATLEKOLITELTRTNELLKSKFSENGQKNRILSLMLKRNKRETKM 182
Db 496 SETQSLLEQDIATYITECSSLKRSLEQARMEVSEDDK-----ALQLLHDIREQRK 549
QY 183 RGMMAKQEGEMKQVQTSRLESQKIAQLEKGLVSTIEKEKIDKESKTEKLEYIEBIS 242
Db 550 Q-----EIKEQYQAOVEEMRLMNMNQLDEEDIVSARRSDLYESE-----LRESR 593
QY 243 CASQVQYKLDIAOLENLEKND-----EILSKOSLEENIVILSKQVEDL 290
Db 594 LAA---EEFKRKAPECQHKLLKADQKPEVGEYAKLEKINAEOQL-----KIQEL 641
QY 291 NVKCOLLEKEDHVNRRNHNENLNAEMNLKQKFILOEQEHEKLOQKELQIDSLQOE 350
Db 642 QEK---LEKAERAEERLEKLNHRNEDSSGIRKKLVAEERRHSLENKVRLETMERRE 698
QY 351 KESSSHQKLCSPQEEVMVKEKNLFEELKQTLDELQKLOQKEQAERLVKOLEEAKSR 410
Db 699 NRLKDDIOTK-----SQIQOMADKILELEEKHREAQVSAQHLEVHLKQK 743
QY 411 -----AEELKLEELKKEAELE-----KSSAAHTQATLLLOEK-----YDSMVQSL 453
Db 744 EQHYEEKIKVLDNQIKKDLADKETLENMMQRHEEEAHEKGKILSEQKAMINAMDSKIRSL 803
QY 454 EDVTAQESYKALTASE-----IEDKLNSLSLOEKAAGNAEDVOHQI 499
Db 804 EQRIVELSEANKLAANSLSLFTQRMNAQOEEMISELROQKFYLETQAGKLEAQRNKLQEL 863
QY 500 LATESSNOEYVRMLDLQTKSALKETEIKETVSLQKITDLOLQKQOEDF-----R 553
Db 864 EKISHODSHKSRILLELETRLEVSLEHEEQKLEKRLQTLQSLQRESQLTALQAR 923
QY 554 KOLEDE--EGRKAQKENTAELETEINKWRLLYEELYNKTKPFLQLODAFEVEKQALLNEH 612
Db 924 AALESQRLQAKTELETTAEAEETQALTAHRDEI-----QRKFDALRNSCTVITD-- 974
QY 613 GAAEOELNKIRDSYAKLLGHON--LKOKIKHVVKLKDENSOLKSEVSKLRQOLAKKQSE 670
Db 975 --LEEQLNQLTEDNAE--LNNQNFYLSKOLDEASGANDEIVOLRSEVDHLRREITEREMQL 1031
QY 671 TKLOEEL 677
Db 1032 TSQKQTM 1038

```

```

RESULT 5
US-09-919-172-29
; Sequence 29, Application US/09919172
; Patent No. US20020119463A1
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary
; TITLE OF INVENTION: PROSTATE CANCER MARKERS
; FILE REFERENCE: PA-0036 US
; CURRENT APPLICATION NUMBER: US/09/919,172
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/222,469
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 29
; LENGTH: 2125
; TYPE: PR
; ORGANISM: Homo sapiens
; FEATURE:

```

```
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020119463A1 3774181CD1
US-09-919-172-29

Query Match      10.0%; Score 358.5; DB 10; Length 2125;
Best Local Similarity 21.0%; Pred. No. 1.4e-08;
Matches 179; Conservative 157; Mismatches 268; Indels 249; Gaps 37;

Qy 26 VKTLEVLKGPVSFQKRSQRFKQKESKQNLND-----KDTTLP-----ASA 66
   | : : : : : | | | | | : : : : : | : : : : : | : : : : :
Db 764 LKLIKQMEKDLAFQKQAEKQKLE-KQKIELEARRKITEIQYTCRENALPVPITQATSC 822
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

Qy 67 RKVKSSEKSKSQKNDKDLKILEKIRVLLQERGAQDRRIODLETELEKMEARLNAALRE 126
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 823 RAVTGLQOQEHDKQ-----AELKQQVDELTAANK-----AEQDMRELYELNALQLE 871
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

Qy 127 KTSLSANNATLEKQLIETRTNELLKSF-----SENGNQNLRLSLEMLKLNKRETK 181
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 872 KTSSEKARLLKLDNETNNTLRLCKLELRKQDAEKGYSQLRGRLQNTTGGKAEPA 931
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

Qy 182 MRGMAKQEGMEMK-----LQVTORESOGKIAQLEGLKLVSTKEKIDE-KSETEKLL 235
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 932 M-----QEAADLKIKKNYOLESLNHEKQKQKQ-----EVDKITRAHVAEKNI 977
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

Qy 236 FYI-EISCASDOVEKYKLDIAQ-----LEENLKEKNDLIL-SLKQSLLEN--IVILSKQ 286
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 978 QHLNLSQIHFRDEKELERLQICQKSDHLKQEFKSHQELLONTAKAENNDKIORLNEE 1037
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

Qy 287 VEDLVNKKQLEKKEKEDHVNRRHN-----ENLNAEMQNLKQKFILEQOE-HEKLOQRELO 342
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 1038 LEKSENECAEMLKQKVEELTRONNETKLMQRIQAESENI-----VLEKQTIQORCEALKIQ 1093
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

Qy 343 IDSLLQKEKELSSSLH-----QKLCFSQFQEMVMEKRLFEELKOTLDELDKLOQK 392
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 1094 ADGFKDQLRSTNEHLHKQTKTEQDFQRIKCLEEDLAKSONLV-SEPKQKCDQOQNIQ- 1151
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

Qy 393 EQAERLVQLEEAASRAELKLEELKKEAELEKSS-----AAHTQAT 439
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 1152 ---NTKKYVRLNAELNASKEERKRGQKQVQLQQAQVQELNNRLKVKQVDELHLKTKTEQMT 1209
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

Qy 440 ---LLQEKYDSMVQSLVEDVTAQFESY---KALTASIEDLKLENSLSQKAAKAGNA 492
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 1210 HRKMVLFQOESGKFKQSAEPEFKKMEKLMESKVITENDISGIRLDFVLSQOENSRAQENA 1269
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

Qy 493 EDVQHQILATESSENOEYVRMLDLQTKSALKETEIKETVTSFLOKITDLQNLKQOEEDF 552
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 1270 K-----LCETNIKE-----LERQLQOYREOM 1290
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

Qy 553 R-----KQLEDE-EGRAEKENTTAELTEINKWRLLYEELYNKTKPFQLOL 598
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 1291 QOGQHMEANHYKQCKQKLEDELIAQKREVENLKQMDQOI-----KEHEHQLVL 1338
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

Qy 599 DAFEVEKQALLN-----EHGAAQEQNLKIRDSYAKLLGH----- 632
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 1339 LOCEIQKSTAKDCTFKPDFEMTVKECHSGELSSRN-----TGHLPHTPRSPLLR 1389
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

Qy 633 ---ONLKQKIKH-----VVKLKDENSQKSEVSKRLCOLAKKQSETKLQ-EE 676
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 1390 WTQEPQPLEEKWOHVRVQIPEKVFQPPGAPLEKESQ-QCYSEYFSQSTELQITFDE 1448
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

Qy 677 LNKVLGKIHFD-----PSKAFHESKENFALK--TPLEKGNVNCY----- 714
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 1449 TNPIRFLSEIKIRDOALNNSRPVRYODNACEMELVKVLPLEITAKNKQYDMHTEVTTL 1508
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

Qy 715 ---RAPMECQESW 724
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 1509 KOEKNPVPSPAEW 1521
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

; GENERAL INFORMATION:
; APPLICANT: KAPPELLER-LIBERMANN, Rosana
; TITLE OF INVENTION: 13245, A NO. US20020160483A1el Human Myotonic Dystrophy Type P
; FILE REFERENCE: 10147-57U1
; CURRENT APPLICATION NUMBER: US/10/017,216
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/242,429
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1597
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-017-216-6

Query Match      9.9%; Score 355; DB 9; Length 1597;
Best Local Similarity 21.4%; Pred. No. 1.5e-08;
Matches 137; Conservative 139; Mismatches 262; Indels 102; Gaps 16;

Qy 91 EIRVLLQERGAQDRRIODLETELEKMEARLNAALREKTSLSANNATLEKQLIETRTNEL 150
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 5 EEAEMQEMTRLHRRVSEVEAVLSQKEVELKASQTSRSLLEQDLATYITESSKRSLEQ 64
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

Qy 151 LKSFSENGNQNLRLSLEMLKLNKRETRKMRGMAKQEGMEMKLVQVTORESLESOQKI 210
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 65 ARMEVSOEDDK-----ALQLLHDIREQSRKLQ-----EIKEQEQYQAQVEEMLMM 109
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

Qy 211 AOLGEGKLV-----IEKEKIDEKSETEKLELYIEIISCASDOVEKYKLDIAOLE 259
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 110 NQLEEDLVARRRRSDLYESLRESRLAAEFKRNACQHKLMKAKQDGKPEVGEYSKLE 169
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

Qy 260 ENLKEKNDLILSKQSLLENIVILSKOVEDLVNKCQLE---KEKEDHVNRRHNENLNA 317
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 170 KINAEOQLKIQLEQKLE-----KAVKASTEATELLONIROAKERAERELEKLNRED 222
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

Qy 318 EMQNLKQFILEQHEKHEKQKELQIDSLQOEKELSSSLHQKLCFSQFQEMVMEKKNLFEE 377
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 223 SSEGIKKKLVAEERRHSLNENKVRLETMERRENRLKDDIQTK-----SE 267
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

Qy 378 ELKQTLDELQKQKEQAERLVKQLEEAASR---AEELKLEELKKEAELEK--- 430
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 268 QIQQMAKQKILEKHEKHEAQVSAQHLVHLKQKQEHVEEKIKVLDNQIKKLADKESLEN 327
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

Qy 431 ---SSAAHQATLLQEK-----YDSMVQSLVEDVTAQFESYKALTASE----- 470
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 328 MWQHEEEAHEKQKILSEOKAMINAMDSKIRSLQRIVELSEANKLAANSSLTQRMKA 387
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

Qy 471 ---IEDLKLENSLSQKAAKAGNAEDVQHQILATESSENOEYVRMLDLQTKSALKETE 526
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 388 QEEMISELROQKFVLETQAGKLEAQRNKRLEQLEKISHQDHSKRSRLELETRLSVSLE 447
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

Qy 527 IKETVTFLOKITDLQNLKQOEEDF-----RKQLEDE-EGRAEKENTTAELTEELINK 579
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 448 HEEQKLEKRLQTLQELQSLQERESQTLQALQARAALLESQLRQAQKTELEETAAEEETQA 507
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

Qy 580 WRLLYEELYNKTKPFQLOLDAFEVEKQALLNEHGAQEQNLKIRDSYAKLLGHQHN--LKQ 637
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 508 LTAHREI-----QRKFDALRNSCIVITD-----LEEQLNQLTEDNAE-LNNQNFYLSK 555
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

Qy 638 KIKHVVKLDNSQKSEVSKRLCOLAKKQSETKLQEEEL 677
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 556 QLDEASGANDEIVQLRSEVDHLRREITERENMLTSQKQTM 595
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

RESULT 7
US-09-884-001-19
; Sequence 19, Application US/09884001
; Publication No. US20020182656A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; Peschon, Jacques J.
```


APPLICANT: Sims, John E.
APPLICANT: Virca, G. Duke
APPLICANT: Willis, Cynthia R.
TITLE OF INVENTION: Methods for Regulating Vascularization Using GEF
TITLE OF INVENTION: Containing NEK-Like Kinase (GNK)
FILE REFERENCE: Immunex GNK/sgnk pct
CURRENT APPLICATION NUMBER: US/09/884,001
CURRENT FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: 60/113,003
PRIOR FILING DATE: 1998-12-18
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn ver. 2.0
SEQ ID NO 19
LENGTH: 868
TYPE: PRT
ORGANISM: Homo sapiens
US-09-884-001-19

Query Match 9.9%; Score 354.5; DB 9; Length 868;
Best Local Similarity 22.0%; Pred. No. 9.4e-09;
Matches 190; Conservative 150; Mismatches 293; Indels 229; Gaps 31;

Qy 25 DVKTEVLKGPVSQKSORFKQKESKONLVNDKDTLPASARKVKSESKEKQNDKD 84
Db 21 DLRDQEEKEIQKLSERHQEAAA-----TTQLEQLHQAQKQEEVLARAVQKE 72
Qy 85 LKILEK---EIRVLLOERGAQD--RRIQDLELEKMEARLNAALREKTSLSANNATLEK 139
Db 73 ALVREKAALEVLQAVDRDQDLAQGLSSAKELLESEFFEAQQQSVIEVTKGQLEV 132
Qy 140 QLIETRTNELLKSKFSNGKNLRIILSLMLKRNKRETKMRG-----MMAKQEG--- 191
Db 133 QIQVTQAKEVI-----QGEVRLCKLELDTERSQAQERDAAARQAQAEQEGTKA 183
Qy 192 -----MEMKLOVQORSLEESQ 208
Db 184 LEQKAAHEKVNOLREKWEKERSWQOELAKALESREKMEMLRLKEQOTEMEAIQ- 242
Qy 209 KIALEGKLVSTIEKIDKSETEK-----LLEYTEITSQASDOVEKYKLDI----- 255
Db 243 --AQREERTQASALCOMQLETEKERVSLLETLLQTKELADASQQLERLRQDMKVOKL 300
Qy 256 -----AQLEENLEKKNDEILSLKQ---SLEENIVILSKOVEDLNKVC- 294
Db 301 KEQETTGILQLOQAQRELKAAQRHDDLAALQEESSLLQDKMDIQKQVEDLKSQVL 360
Qy 295 -----QLLEKEKEDHVNREHN---ENLNAEQNL-----TKQFLEQOEHEKLOQK 339
Db 361 AQDQSRLVEQVEQKURETOEYNRIQKELERERKASLTLSMERKEORLLVLEQADSIHQ 420
Qy 340 ELQIDSLQOEKELSSSLHQKLCFOERM---VKEKNL-FEELKQTLDELKLOQKEE 394
Db 421 EL---SALRQDMQBAQ--EQKELSAQWELLQVEKKEADFLAQEAQLLELEASHITEQ 476
Qy 395 QAERLVKOLEEAKSRABELKLEELKKEKAELEKSSAAHTQATL-----LLQEKYDSM 449
Db 477 QLRASLWAEAKAAQQLRLRSTESQLEALAAEQPGNQAQQAQALASLYSALQOALGSV 536
Qy 450 VQSLDVTQAQTES-----YKALPASEI-----EDL----- 474
Db 537 CESRPELGGGDSAPSVNGLPEPDONGARSLFKRGPPLLTALSAAVASALHKLHODLMWTK 596
Qy 475 -----KLENSISLQEAQAAGKNAE--DYHQHILATESSNOEYVRLMLDLOTSA 521
Db 597 QTRDLVRDQVOKLERLTDTBAEKSQVHTELQDLQORQL-----SONQEE-----KSKWE 645
Qy 522 LKTEIKETVTSFKQITDQNLQKQDEDFRKOLEDDEGRKAENKTTA----- 571
Db 646 GKQNSLESELMELHETWASLQSRRLRAELQRMQAQGERELLQAAKENLTAQVEHLQRAVV 705
Qy 572 ELTEINKRWRLLYELYNKTKPFOLOLDAFEVEKQALLNEHGAQOEQLNKTRDSYAKLLG 631
Db 706 EARAQASAAAGILEEDL--RTARSALKLNKEEVESE---RERAQALQOEQELKVAQGRALQ 760

Qy 632 H-----ONLKQIKHVVVKLDKENSOLKS--EVSKLRCQL-----AKKKOSETKIQ-----E 675
Db 761 ENLALLTQTLAREEEVETLRGQIOELEKQREMQRAALELLSLDLKKRNQEVDLQOEQIQ 820
Qy 676 ELNKVLGIKHFDPSPKAFHESK 697
Db 821 ELEKCRSVLEHLPLMAVQEREQK 842

RESULT 8
US-09-745-763-168
Sequence 168, Application US/09745763
Patent No. US20020065394A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
McCoy, John M.
LaVallie, Edward R.
Collins-Racie, Lisa A.
Evans, Cheryl
Merberg, David
Treacy, Maurice
Spaulding, Vikki
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
ENCODING THEM
NUMBER OF SEQUENCES: 219
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/745,763
FILING DATE: 18-Jun-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 168:
SEQUENCE CHARACTERISTICS:
LENGTH: 677 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 168:
US-09-745-763-168

Query Match 9.8%; Score 353.5; DB 10; Length 677;
Best Local Similarity 25.0%; Pred. No. 8.4e-09;
Matches 175; Conservative 126; Mismatches 242; Indels 157; Gaps 33;

Qy 69 VKSESKEKESQK-----NDRDLKILEK-----EIRVLQERGAQDRRIQDLETEL--- 113
Db 8 LKVSDEMKKSHDAIIDDLNRKLLDVTQYTKKLEMEKLLLENDLSLSDVSRLFTVFPVP 67
Qy 114 EKMEARLNAALREKTSLSANNATLEKQIETLR-----TNEILKSKFSENGKNRILSL 169
Db 68 EKHE-----KEITALSKSNIVELKQLSELKKKCGEQEKIHALTSENTNLK--KMMSN 118
Qy 170 ELMKLRNKRKTKM--RGMMAK-----OEGMEM--KLOVQORSLEESQK 209
Db 119 QYVPVKTHEEVKMTLNDTLAKTNRELLDVKKKFEDINQEFVKIKDKNEILKRNLENTQ 178

QY 210 I-----AQEGKLVSIK--EKIDKSETEKLLLEY-----IBEISCASQVQVEKYKIDIAQ 257
 Db 179 IKAEXISLAHEAKKSSLSQSMRKVOD-SNAELANRYRGQEEIVTLHAEIKRAQKKELD 237
 QY 258 LEENLKEKNDLILKQSLLEENIVILSKQVEDLNKVCOLLEKEK-----EDHVNRRHN 312
 Db 238 IQEIKVKYAPIVSFE--CERFKATEKELKD-----QLEQTKYSVSEEEVKNKQBN 291
 QY 313 ENLNAM-----ONLQKFTLEQOEHEKLOQKELQIDSLQOEKELSSSLHOKLCSFOEM 368
 Db 292 DKLKKEIFLQDLDRKTVLIEKSHMERALSRRKTDENKQKLDLS---QKYTEVKN-- 345
 QY 369 VKEKNLFEELKOTDEL--DKLOQK-----EQABRLVKQLEEEAKSRAEELK----- 415
 Db 346 VREK-LVEENAKQTSILAVNQLQKHVPLEQVEALKSLNGTIENTLKEELKSMQRCVE 404
 QY 416 -----LLE-----EKLKGEA-----ELEKSSAAHTQATLL 441
 Db 405 KQQTQVTKLHQLLENQKSSVPLAEHLQIKAEFEKVEGVIKASLREKEEESQNMKEYSK 464
 QY 442 LOEKYDSMVQSLDV----TAQFESYKALTASEIEDLKLSSLOEKAAGKAGKNAEDVQH 497
 Db 465 LOSEVQNTQALKKLTREVDLSYKATKS-----DLEQISSLNEKLANLRKYEVECE 520
 QY 498 QILATE-----SSNQEVYRMLDLQTSALKETEIKETIVSFQKITDQLNQLKQOEEEDFR 553
 Db 521 EYLHAKKEISAKDE--KELLHFSIEQEIQD--KERCDKSLTTITELQRRIOES----A 572
 QY 554 KQLEDEEGKAKENTTAELTEINKRWLLYELNKTYPFQLOLDAFEVEKQALLNEHG 613
 Db 573 KQIEAKDNKITLNDVRLKQALNGLSLQTYTSGNPTKRQSLDITLQHOVKSLEQQLA 632
 QY 614 AAOEOLNKRDSYAKLL-----GH--ONLKQIKHVVKLK 646
 Db 633 DADRQHQEVIARYTHLLSAAOGHMDVQOEALLQIQMR 672

RESULT 9
 US-10-017-216-5
 ; Sequence 5, Application US/10017216
 ; Patent No. US20020160483A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KAPPELLER-LIBERMANN, Rosana
 ; TITLE OF INVENTION: 13245, A No. US20020160483A1 Human Myotonic Dystrophy Type Pro
 ; FILE REFERENCE: 10147-5701
 ; CURRENT APPLICATION NUMBER: US/10/017,216
 ; PRIOR FILING DATE: 2001-10-23
 ; PRIOR APPLICATION NUMBER: US 60/242,429
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 5
 ; LENGTH: 1641
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-10-017-216-5
 Query Match 9.7%; Score 350; DB 9; Length 1641;
 Best Local Similarity 22.13%; Pred. No. 2.6e-08;
 Matches 158; Conservative 140; Mismatches 285; Indels 126; Gaps 23;
 QY 11 FNDP-----SGCAPSPGADVTKTLVKGVPVFSQKRSQKQESKQNLNVKDTTL 62
 Db 15 FDEPKNSVSSVQCLSPSGFSGELPFV--GFYSYKALGYLGRSES-----VVSIDS 67
 QY 63 PASARKVSKESK-----KESQNKDKDLKLEKEIRVLQERGAQDRRIQDLETELEKME 117
 Db 68 PA---KVSSMEKKLLTKSKELQDSQDKCKMKEQETRL-----HRRVSEAVLSQKE 117
 QY 118 ARLNALREKTSLSANNATKLELQLELRTNELLKSKSENGNQNLRLSLLEMLKLNK 177

Db 118 VELKASQTORSLEQDLATYITFCSSLRKSLRQARMEVSEQEDDK-----ALQLLHDIRE 171
 QY 178 RETMRGMMAQEGMEKMQVQTSRLESOGKIAOLEGKLVSTIEKEIKDEKSETEKLLLEY 237
 Db 172 QSRKIQ-----EIKEQEQYQAVQEMRLMNMNLEEDLVSAARRSDLYESE----- 215
 QY 238 IBEISCASQVQVEKYKIDIAQLEENLKEKNDLILKQSLLEENIVILSKQVEDLNKVCOLL 297
 Db 216 -----LRESLAAAEFKRANECQHKLMKADQCKPVEGYSK--LEKINAEQQLK 264
 QY 298 EKEKEDHVNRRNHNENLNAEMQNLKQKFIHQOEHEKLOQKELQIDSLQOEKELSSSL 357
 Db 265 IQELQEKLEKAVKASTATELQNTQAKERAERLEKLNRE-----DSSEGI 313
 QY 358 HQKLCFSQOEEMVKEKNLFEELKQTLDEL--KLOQKEQAEARLVKQLEEEAKSRAEELK 416
 Db 314 KKLVEAE--LEEKH---REAQVSAQHLVHLKQEHYEEKIKVLDNQIKKDLAKRES 368
 QY 417 LEEKLKGAEAELEKSSAAHTQATLLQEK-----YDSMVQSLDVTYQAFESYKALTASE- 470
 Db 369 LENMQRHEE-----AHEKGKILSEQKAMINAMDSKIRSLERIVELSEANKLAANSS 422
 QY 471 -----IEDLKLSSLOEKAAGKAGKNAEDVQHQILATLATESSENOEVYRMLDLQ 517
 Db 423 LFTORMKAQEMISELROQKPYLETOAQKLEAQNRLKEEQLEKISHQDHSKSLLELE 482
 QY 518 TKSALKETEIKETIVSFQKITDQLNQLKQOEEDE-----RKQLEDE--EGRKAKENTT 570
 Db 483 TPLREVSLEHEEOKLELKRQLTELQLSQERESQLTALQARAALLESQLRQAKTELEETT 542
 QY 571 AELTEEINKRWLLYELNKTYPFQLOLDAFEVEKQALLNEHGAQOQLNKRDSYAKLL 630
 Db 543 AEABEEIQALTAHRDEI-----QRKFDALNCSVTITD-----LEEQLNQLTDEANAE-L 590
 QY 631 GHON--LQKIKHVVKLKDENSQKSEVSKRCLQAKKKQSETKLQEEEL 677
 Db 591 NNQNFYLSQKDEASGANDEIVOLRSEVDHLRREITEREMQLTSSQKQTM 639
 RESULT 10
 US-09-864-761-47959
 ; Sequence 47959, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
 ; FILE REFERENCE: Aecomica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30

```

; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47959
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AJ010770.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
; OTHER INFORMATION: SWISSPROT HIT: Q99323, EVALU6 3.00e-17
; OTHER INFORMATION: EST_HUMAN HIT: AU132932.1, EVALU6 1.00e-10
; US-09-864-761-47959

```

Query Match	9.08;	Score 323;	DB 10;	Length 660;		
Best Local Similarity	22.5%;	Pred. No. 1.6e-07;				
Matches 161; Conservative 135; Mismatches 238;	Indels 180;	Gaps 30;				
QY	67	RKYSSE----	SKKESQKNDKDLKILEKIRVLLQERGAQDRRIQDLFTELEKWEARLNA	122		
DB		: :	: :	: :		
QY	4	KVENSKEEIQEKETIEBELNTKIIIEBKYL-----	ELKDKLTTADKLLGELQE	54		
DB		: :	: :	: :		
QY	123	ALREKTSLSANNATLEKOLIELTRTNELLKSKFSENGKNLRLI--	SLEMLKURNKRET	180		
DB		: :	: :	: :		
QY	55	QIVQK-----	NQBIKNKMKLELT-----	NSKQERQSSEEIKOLMGTVTEELQKRNHKS	102	
DB		: :	: :	: :		
QY	181	KRGMMAKQECMEMKLQVTPORSLEESQKIAQLSGK-----	LVYSIEKEKIDE	227		
DB		: :	: :	: :		
QY	103	QFETDIVOR--ME---	QETQRKLSQRAELDEMTCQOIVQMKEELIROHMAQMEEMKTRH	157		
DB		: :	: :	: :		
QY	228	KSETKEKLEIETEISCASDQVEKYKLDIAQLEENLKKNDEILSKQSLSEENIVITLSKQV	287			
DB		: :	: :	: :		
QY	158	KGEMENALRSYSNITVEDQIKLNVAINELNKLQDTN-----	SQEKLEKEELGLI----	209		
DB		: :	: :	: :		
QY	288	EDLVKQCLEKEKEDHVNRNREINENLNAEQNLKOKFILEQOEHEKLOQKE---	LOID	344		
DB		: :	: :	: :		
QY	210	--LEEKC-ALQORQLEDVVEELSFSEQIQARQATIAEQESKLNPAHKSLSVTEDLKAV	266			
DB		: :	: :	: :		
QY	345	SLLOQEKELSSSLHOKLCSFO--	EEWKKNLFEELKQFLD-ELDKLOOK-----	BEQ	395	
DB		: :	: :	: :		
QY	267	SASERLELKHPEAVTNYKIKLEMLEKEKANVLDRMAESQAELEKRLTQLLFSHEE	326			
DB		: :	: :	: :		
QY	396	AERLVKOLEEAKSRABELK-----	LLEEKLUKGEAEL-----	EKSSAAHTQATLL	441	
DB		: :	: :	: :		
QY	327	LSKLKLEDEHTRINIEKLDNLGIHYKQIQDGLQNEMSQKTIETMQPEKDNLTIKQNL	386			
DB		: :	: :	: :		
QY	442	LO-EKYDSMVQSL-----	EDVTAQFESYKALTASEIEDLKLNSLSQKAKAKGNAEDV	495		
DB		: :	: :	: :		
QY	387	LETSKLKDQOQSLVNSKEEMTLQ-----	INELQKEIETLROEKEKGTLEQEV	435		
DB		: :	: :	: :		
QY	496	QHQILATESSNQEVYRMLLDIQTKSALKETBIKETITVSFLQITDLOLQKQOE-----	549			
DB		: :	: :	: :		
QY	436	Q-----	ELQKTELLEKQMK-----	KENDLOEKFPAQLEANSIL	470	
DB		: :	: :	: :		
QY	550	EDPRKQLED-----	EGR-----	KAEKNTTAALTEETI-----	NKWRLLY	584
DB		: :	: :	: :	: :	
QY	471	KDEKTTLEMLKIHTPVSQEERLFILODSIKSKSDVYWEKEIEITILEENEDLKQOCIOLN	530			
DB		: :	: :	: :	: :	

Qy	585	EELYNKTPQLOLDAFEVEKQALLNEHGAAQOLNKIRDSYAKLICHQNLQOKIEHVK	644
Db	531	EEIEKQRTSFBAKNEPVNYQLQEYAC-----LUKDD-----LEDSKNQOEYSKSK	592
Qy	645	LKDNSQLKSEVSKRLQOLAKKQSETKLOEELNKVLGINKHFDPSKAFHESKE	698
Db	583	LKALNEEL--HLORINTPTTKMKSSV-----FDEKTFVAETLE	619

RESULT 11

```

US-09-815-242-10617
; Sequence 10617, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA-011A
; CURRENT APPLICATION NUMBER: US/09/815,
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-03-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSQ for Windows Version
; SEQ ID NO 10617
; LENGTH: 1045
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10617

```

Query Match	8.9%	Score 320;	DB 10;	Length 1045;
Best Local Similarity	22.3%;	Pred. No. 3.2e-07;		
Matches	155;	Conservative 140;	Mismatches 223;	Gaps 31;
Qy	38	FOKSOREK-QOKESKONLVNVDK---TTILPASARKVKSSKESKESQKNDKDLKILEKEIRV	94	
Db	230	FEKQLLESQOQEAQQLLVREQIATL-----KQAKQAKEQVRY	269	
Qy	95	LLQERGAQDRRIQDLETELEKMEARLNAALREKTSLSANNATLEKOLIELTRTNELLSK	154	
Db	270	AIER-----QNLQOQKE-----LLEKAKAQEQETVIERLKE-----OIOLK	309	
Qy	155	FSENGKNRLRILSLMLKLNKRET---KMRGMMAKOBGEMKMLQVTRLSLEESQGIQAQ	212	
Db	310	WSQKQOSLAEKVFKEKSEKQOKEQETYSKQALMETQOAL---TDWQAINSELEEQOPLAE	368	
Qy	213	LEGKLVSIEKEKIDEKSETEKLLRYEIEISCAQVEK---YKLDIAQLLENLKE---N	266	
Db	369	KQERLQITQIQ-----LPQYQEYQLAQQAQIAEQANYQAIQKEYESCQOKEKITLA	418	
Qy	267	DEILSKOSLEENIVILSKQVEDLVNVCOLLEKEKEKHVNRRHNE-----NUNAMQ	320	
Db	419	DKVATAKQIFQO-----EGTLEKANFECSSVADHWONFEVERKONKAWOKISONOVELH	473	

Qy	78	SOKNDXDKLILEKIRVILQERGAQDRRQDLETELEKKEARLNAA	-----LRE 126
Db	38	SLUEDK--IRLLQDLEVERELRIERAKADISVOIQMSERLEEAGGAHQFEANRKR	96
Qy	127	KTSLSANNATLEKQELIETRTNELLSKTFSE--NGOKNRLTSLLEMLKRLNKRKTKMRG	184
Db	97	DAELLKLRLLEDVHLESEETTLLAKKHNEIITDQEQVEIILT-----KKN-----	143
Qy	185	MMAKQEGEMKMLQVTORSLEESOGKTAQLEGKLIVSTEKEKIDEKSTEKLBYIEISCA	244
Db	144	--ARAEDKAKFOT-----EYVELLSQTESYNKEKIVSEKHISLEVISISELNVK	191
Qy	245	SDQVEKYKIDAOELNEKKNDEIILSLQAOLEENIVILSKOVEDLVNKKCOLLEKEDH	304
Db	192	IEELNRTVIDISHSRSLSGENTE-----UTKDQDDLUKVDTVSFSKSOV	237
Qy	305	VNRNREHNENLNAEMONLQKOFLEOQEHKELOOKELQIDSL---LQOEKELSSLDHQKL	361
Db	238	ISOLEDARRRLEDED---RRRSLESSLH---QVEIELDSVRNQLSESEARIDLERQL	290
Qy	362	CSQOEEMVKEKNLFEBELKOTLDLDELKLOOK-----EQOARLVKOLE--EAKSR-	410
Db	291	VKANADLTSQNKWNGSEVAAAREVEIERRKYQVRIETEELHEHIESLIVKNVKNLEKMKTRL	350
Qy	411	AELKLK---EELKLGKEALEKS-----SAAHTQATILLQOE-----KYDSM	449
Db	351	ASEVEVLIIDLEKSNNSCRELTKSVENTLEKHNVELASRLDETIILYETQSOLDKKNKHADL	410
Qy	450	VQSLDEDVTAQFESYKALTAS-----EIEDLKLENS-----S	480
Db	411	VRTVHELDKYKDNNNTOTRENKLGDDLHEAKGAINELNRRLLHELEURLRLENERDELT	470
Qy	481	LQKAAKAGNAEDVOHQIILATESNOQVYRMLLDIQTQSALKETEI---KEITVFSFQ	536
Db	471	ANYKEAEAGRAKEEQRCRLAADFNQYRH-----DAERRLAEKDEIEIAIRKQTSIETQ	525
Qy	537	---KITDLOQLKOQBEDFRKQLE---DEEGRAKAKENTTALTEETINKWRLLYEELYN	589
Db	526	LNARVIEAETRLKTEVTRIKKLIQIQTILEMSLDVANKTNDLQVKVKKQSILQTLTQQA	585
Qy	590	KTAPFQOLDFAFEVEKQALLNEHGAEOOLNKIRDSYAKLLGHONLKQKIKHVVKLKDSN	649
Db	586	HYEDVQRQL-----QATLDQYAVAQRRLAGLNGELEEVRSHLDSANRAKRTVVELOYE-	637
Qy	650	SOLKSEYSKLRCOLAKKQKSETKLOBELNKK	680
Db	638	-EAASRLNELLTANVSLVSTKLEOELSV	667

RESULT 14

```

US-09-815-242-13262
; Sequence 13262, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26

```

[illegible]

RESULT 15
US-09-815-242-13608
; Sequence 13608, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13608
; LENGTH: 1179
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13608

Query Match 8.6%; Score 309; DB 10; Length 1179;
Best Local Similarity 21.5%; Pred. No. le-06;
Matches 166; Conservative 149; Mismatches 267; Indels 190; Gaps 29;
Qy 68 KVKSESKESKOKNDKLEKEIRVLQERGAQDRRIQDLELEKWEARLNAALREK 127
Db 171 KTRRKETESKLOQTQDNLDRLDIYEL-----DNQIKPLEKQAEARFLEGGORK 223
Qy 128 -----TSLSANWATLEKQILTRTNELLKSKFSENGNKNLRILSLMLKLRNKR 179
Db 224 ATYLDVLVAQIKDKNAELDSTEELAQVOELLTSYV-----QKREK-LEENQTLKKQRQ 277
Qy 180 TKMRGCMMAKQEGMEM-----KLQVTRSL--EESQKIAQLEGKLVISIE 221
Db 278 -DLQAEAMKQDQSLMDLTSLSLDERKLALSKEQVAINQOEQAARLATLEDKRSSL 336
Qy 222 KEKIDKSETEKLLLEYIEBISCASQVBEKLDIAOLEENLEKRNDEILSLK----- 273
Db 337 QEKSKES-----SLALLEGNLVQNNQKLRNLEAEALLAFSD 372
Qy 274 -----QSLREENIVILSKOVEDLVNKKCOLLEKEKEDHVNRRNHNENNAEMQNLKQKFI 327
Db 373 DPQMIELLRRFVALLQEEADVSNQLFRINELN-----SRQLSQKQADQLEKLKEQLA 428
Qy 328 LEQQEHEKLOQKELQIDSLLOEKELSSHLKCLCSFOEEMVKEKNLFEELKQTLDEL 387
Db 429 TAK-----EKASQOKEELETAKEQVOKLLAD-YQATAKEQEE---QKTSYQAQOQSLFRLD 481
Qy 388 KLOQKEQAERLVKOLEEAKSRABEKLLEK-----LKGKEAE-----LEKS 431
Db 482 NLKNQARAQSLNLRNHSNFYAGVKSVLQEKDRLGGLIIGAVSEHLTFDMHYQTALeia 541
Qy 432 SAAHTQATLLQOE-----KYDSMVQSLSDVTAQFESYKA 465

Db 542 LGASSOHLIVEDENAAATKAIDFLKRNAGRATFLPLTTIKARTISSONQDAIAVSPGLG 601
Qy 466 LTASEIE--DLKLE-----NSSLOEKAAGKNAEDVOHQIL----- 500
Db 602 M-ADELVTFTDRLEAIFKNLLATTAFDTVEHAREAAQVRYQVRWVTLDTGTELTGSGY 660
Qy 501 ---ATESNQBYVRMLLD-LQTKSALKETEI--KEITVSFLQ-----KITD-LQNLKQOEE 550
Db 661 AGGANRQNSIFIKPELEQLQKEIAEAEASLDSEVALKTQDEMARLUTESLEAIKSGE 720
Qy 551 DFRKOLEDEEGRKAEKENTTAELTEEINKWRLLLYEELYNKT-----KPFQLQDAFE 602
Db 721 QARIQ---EOGLFLAYQOOTSQOVEELETMLKLBEEIDRLSEGDWQADKENCQERLAAIA 777
Qy 603 VEKQALLNEHGAQOEQLNKIRDSYAKLLGHQNLKOKIKHVVKLKD-NSQLKSEVSKLRC 661
Db 778 SKONLEAEIEEIKSNKNAIQERY-----ONLQELAAQARLLKTELOGQKRYEVADIE- 830
Qy 662 QLAKKKQSETKLQEEELNKLGIKHFDPSKA----FHESKENFALKTPLEK 709
Db 831 RLKELDNLTQEIEIQRMLOEKVDNLEKVDTLNNOQAEEAKTKTNLQOG 882

Search completed: December 20, 2002, 15:14:40
Job time : 16.0892 secs


```
QY 224 KIDKSETEKLELEYIEISCDQVEKYKIDIAQLEENLKEKNDILSLKQSEENIVL 283
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 KIDKSETEKLELEYIEISCDQVEKYKIDIAQLEENLKEKNDILSLKQSEENIT-F 119
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 284 SKQVEDLVKCOLLEKEKEDHVNRRHNENLNAEQNLKQKFIQEQHEKLOQKELQI 343
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 120 SKQIEDLVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALERQVEYKLOQKELQ 179
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 344 DSLLOQEKELSSSLHOKLCSFOEEMVKNLFEELKQTLDELKLOOKEQEAERLVKOL 403
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 180 QSLLOQEKELSSARLOQLCSFOEEMTSEKNVFEELKALAEALDAVQOKEQSERLVKOL 239
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 404 EBEAKSRAEELKLEKLGKAELEKSSAAHTQATLL----- 441
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 240 EBEKSTAEQLTRLDNLREKEVELEKHTAAHAQAAILIAQEKYNDTAQSLRDVTAQLESV 299
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 442 ----- 441
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 300 QEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQL 359
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 442 --LOEKYDSMWQSLQEDVTAQFESYKALTASETEDLKLENSLSQEKAAKAGNAEDVOHQI 499
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 360 ESVQEKYNDTAQSLRDVTAQLESYKSSLTKEIEDLKLENLTLOEKVMAEKSVEDVQOOI 419
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 500 LATESNQYVRLMLDLQTKSALKETEITEITVSFLQKITDLQNLKQOQEEFRKQLEDE 559
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 420 LTAESTNQYARMVDLQNRSTLKEEKEITSSFEKITDLKQNLQOQDEFRKQLEBK 479
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 560 EGRKAENNTAEITAEINKNRLLYEELYNKTKPFQQLDAFEVEKQALLNEHGAQEOQL 619
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 480 GKRTAEKENVMTELTMEINKNRLLYDELYEKTTPFQOQLDAFEKQALLNEHGAQEOQL 539
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 620 NKIRDSYAKLLGHQNLKQIKHVVVKLDKNSQLKSEVSKLRQAKKQKQSETKLOEELNK 679
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 540 NKIRDSYAKLLGHQNLKQIKHVVVKLDKNSQLKSEVSKLRQAKKQKQSETKLOEELNK 679
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 680 VLGKHFDPSPKAFHHSKENFALKTPLEKGNPNC 713
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 600 ALGIRHDFPSKAFCHASKENF---TPLKEGNPNC 630
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 2
US-08-477-831C-2
; Sequence 2, Application US/08477831C
; Patent No. 6429291
; GENERAL INFORMATION:
; APPLICANT: TURLEY, EVA A.
; APPLICANT: SHUWEN, ZHANG
; APPLICANT: ENTWISTLE, JOYCELYN
; TITLE OF INVENTION: HVALURONAN RECEPTOR PROTEIN
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10020-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Rel. #1.0, ASCII
; CURRENT APPLICATION DATA: US/08/477.831C
; APPLICATION NUMBER: US/08/477.831C
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: PIERRE, MARGARET A.
; REGISTRATION NUMBER: 30,709
; REFERENCE/DOCKET NUMBER: SIM-10
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 606 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: /desc = "RHAMM I protein"
US-08-477-831C-2

Query Match 50.6%; Score 1819.5; DB 4; Length 606;
Best Local Similarity 61.0%; Pred. No. 3.4e-95;
Matches 387; Conservative 50; Mismatches 84; Indels 113; Gaps 4;

QY 164 LRLISLELMKLNKRETKMRGMMAKQEGNEMKLVQTSRLESQGIKIAQLEKLVYSIEKE 223
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MRALSLELMKLNKRETKMRGMMAKQEGNEMKLVQTSRLESQGIKIAQLEKLVYSIEKE 54
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 224 KIDKSETEKLELEYIEISCDQVEKYKIDIAQLEENLKEKNDILSLKQSEENIVL 283
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 55 -----CASDQVEKCKVDIAQLEEDLKEKDRILSLKQSEENIT-F 94
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 284 SKQVEDLVKCOLLEKEKEDHVNRRHNENLNAEQNLKQKFIQEQHEKLOQKELQI 343
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 95 SKQIEDLVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALERQVEYKLOQKELQ 154
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 344 DSLLOQEKELSSSLHOKLCSFOEEMVKNLFEELKQTLDELKLOOKEQEAERLVKOL 403
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 155 QSLLOQEKELSSARLOQLCSFOEEMTSEKNVFEELKALAEALDAVQOKEQSERLVKOL 214
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 404 EBEAKSRAEELKLEKLGKAELEKSSAAHTQATLL----- 441
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 215 EBEKSTAEQLTRLDNLREKEVELEKHTAAHAQAAILIAQEKYNDTAQSLRDVTAQLESV 274
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 442 ----- 441
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 275 QEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQL 334
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 442 --LOEKYDSMWQSLQEDVTAQFESYKALTASETEDLKLENSLSQEKAAKAGNAEDVOHQI 499
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 335 ESVQEKYNDTAQSLRDVTAQLESYKSSLTKEIEDLKLENLTLOEKVMAEKSVEDVQOOI 394
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 500 LATESNQYVRLMLDLQTKSALKETEITEITVSFLQKITDLQNLKQOQEEFRKQLEDE 559
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 395 LTAESTNQYARMVDLQNRSTLKEEKEITSSFEKITDLKQNLQOQDEFRKQLEBK 454
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 560 EGRKAENNTAEITAEINKNRLLYEELYNKTKPFQQLDAFEVEKQALLNEHGAQEOQL 619
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 455 GKRTAEKENVMTELTMEINKNRLLYDELYEKTTPFQOQLDAFEKQALLNEHGAQEOQL 514
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 620 NKIRDSYAKLLGHQNLKQIKHVVVKLDKNSQLKSEVSKLRQAKKQKQSETKLOEELNK 679
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 515 NKIRDSYAKLLGHQNLKQIKHVVVKLDKNSQLKSEVSKLRQAKKQKQSETKLOEELNK 679
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 680 VLGKHFDPSPKAFHHSKENFALKTPLEKGNPNC 713
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 575 ALGIRHDFPSKAFCHASKENF---TPLKEGNPNC 605
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
US-08-402-217A-2
; Sequence 2, Application US/08402217A
; Patent No. 5587301
; GENERAL INFORMATION:
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Wilge, Craig G.
; APPLICANT: Seilhamer, Jeffrey J.
; TITLE OF INVENTION: HVALURONAN RECEPTOR EXPRESSED IN HUMAN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
```


ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/402,217A
FILING DATE: 10-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Luther, Barbara J.
REGISTRATION NUMBER: 33954
REFERENCE/DOCKET NUMBER: PF-0028US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-852-0195
TELEFAX: 415-852-0195
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 351 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-402-217A-2

Query Match 47.1%; Score 1694; DB 1; Length 351;
Best Local Similarity 98.9%; Pred. No. 2.2e-88;
Matches 345; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 319 MNLKQKFIHQEQHEKQLQKELQIDSLQKELSSSLHKLCSFQEMVKEKNLFEE 378
Db 1 MNLKQKFIHQEQHEKQLQKELQIDSLQKELSSSLHKLCSFQEMVKEKNLFEE 60
Qy 379 LKQTLDELKLOQKEQAEERLVKQLEERAKSRAEELKLEELKKEAELEKSSAAHTQA 438
Db 61 LKQTLDELKLOQKEQAEERLVKQLEERAKSRAEELKLEELKKEAELEKSSAAHTQA 120
Qy 439 TLLQEKYDSMVQSLQEDVTAQFESYKALTASEIEDLKLENSSLQEKAAKAGNAEDVQHQ 498
Db 121 TLLLEEKYDSMVQSLQEDVTAQFESYKALTASEIEDLKLENSSLQEKAAKAGNAEDVQHQ 180
Qy 499 ILATSSNQEVYRMLDLQTKSALKETEIKETVTSFLQKITDLQNLKQOEEDFRKQLED 558
Db 181 ILATSSNQEVYRMLDLQTKSALKETEIKETVTSFLQKITDLQNLKQOEEDFRKQLED 240
Qy 559 BEGRKAKEKNTAETELTEINKWRLLYEELYNKTPFQQLDAFEVEKQALLNEHGAQEQ 618
Db 241 BEGRKAKEKNTAETELTEINKWRLLYEELYNKTPFQQLDAFEVEKQALLNEHGAQEQ 300
Qy 619 LNKIRDSYAKLGHONLKQIKHVYKVKLDKNSQLKSEVSKLRCLAKKK 667
Db 301 LNKIRDSYAKLGHONLKQIKHVYKVKLDKNSQLKSEVSKLRCLAKKK 349

RESULT 4
US-08-700-178-2
; Sequence 2, Application US/08700178
; Patent No. 5783669
; Patent No. 5783669 5700912
; GENERAL INFORMATION:
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Wilde, Craig G.
; APPLICANT: Seilhamer, Jeffrey J.
; TITLE OF INVENTION: HVALURONAN RECEPTOR EXPRESSED IN HUMAN
; TITLE OF INVENTION: UMBILICAL VEIN ENDOTHELIAL CELLS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/700,178
FILING DATE: August 20, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/402,217
FILING DATE: March 10, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0028-1 DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 351 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-700-178-2

Query Match 47.1%; Score 1694; DB 1; Length 351;
Best Local Similarity 98.9%; Pred. No. 2.2e-88;
Matches 345; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 319 MNLKQKFIHQEQHEKQLQKELQIDSLQKELSSSLHKLCSFQEMVKEKNLFEE 378
Db 1 MNLKQKFIHQEQHEKQLQKELQIDSLQKELSSSLHKLCSFQEMVKEKNLFEE 60
Qy 379 LKQTLDELKLOQKEQAEERLVKQLEERAKSRAEELKLEELKKEAELEKSSAAHTQA 438
Db 61 LKQTLDELKLOQKEQAEERLVKQLEERAKSRAEELKLEELKKEAELEKSSAAHTQA 120
Qy 439 TLLQEKYDSMVQSLQEDVTAQFESYKALTASEIEDLKLENSSLQEKAAKAGNAEDVQHQ 498
Db 121 TLLLEEKYDSMVQSLQEDVTAQFESYKALTASEIEDLKLENSSLQEKAAKAGNAEDVQHQ 180
Qy 499 ILATSSNQEVYRMLDLQTKSALKETEIKETVTSFLQKITDLQNLKQOEEDFRKQLED 558
Db 181 ILATSSNQEVYRMLDLQTKSALKETEIKETVTSFLQKITDLQNLKQOEEDFRKQLED 240
Qy 559 BEGRKAKEKNTAETELTEINKWRLLYEELYNKTPFQQLDAFEVEKQALLNEHGAQEQ 618
Db 241 BEGRKAKEKNTAETELTEINKWRLLYEELYNKTPFQQLDAFEVEKQALLNEHGAQEQ 300
Qy 619 LNKIRDSYAKLGHONLKQIKHVYKVKLDKNSQLKSEVSKLRCLAKKK 667
Db 301 LNKIRDSYAKLGHONLKQIKHVYKVKLDKNSQLKSEVSKLRCLAKKK 349

RESULT 5
US-08-995-654-2
; Sequence 2, Application US/08995654
; Patent No. 6025138
; GENERAL INFORMATION:
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Wilde, Craig
; APPLICANT: Seilhamer, Jeffrey
; TITLE OF INVENTION: HVALURONAN RECEPTOR EXPRESSED IN HUMAN
; TITLE OF INVENTION: UMBILICAL VEIN ENDOTHELIAL CELLS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/995,654
FILING DATE: December 22, 1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/700,178
FILING DATE: August 20, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/402,217
FILING DATE: March 10, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0028-2 DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 351 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-995-654-2

Query Match 47.1%; Score 1694; DB 3; Length 351;
Best Local Similarity 98.9%; Pred. No. 2,2e-88;
Matches 345; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 319 MONLKQKFTLEQOEHEKLOQKELQIDSLLOQKEKLSLSLHQKLCFSQOEVMKEKNLFEE 378
Db 1 MONLKQKFTLEQOEHEKLOQKELQIDSLLOQKEKLSLSLHQKLCFSQOEVMKEKNLFEE 60
Qy 379 LKQTLDELKLOQKEQAEERLVKQLEEEAKSRAEBLKLEELKLGKGAELKSAHAHTQA 438
Db 61 LKQTLDELKLOQKEQAEERLVKQLEEEAKSRAEBLKLEELKLGKGAELKSAHAHTQA 120
Qy 439 TLLLOEKYDSMVOSLEDVTAQESYKALTASIEDLKLENSLQEKAAKAGNAEDVQHQ 498
Db 121 TLLLEKYDSMVOSLEDVTAQESYKALTASIEDLKLENSLQEKVAKAGNAEDVQHQ 180
Qy 499 ILATSSNQYVYMLDLQTKSALKETEITVTSFLQKITDLOLQKQOEEDFRKQLED 558
Db 181 ILATSSNQYVYMLDLQTKSALKETEITVTSFLQKITDLOLQKQOEEDFRKQLED 240
Qy 559 EGRKAENTTAETETINKWRLLYEELYNKTPQLOLDAFEVEKQALLNEHGAQEQ 618
Db 241 EGRKAENTTAETETINKWRLLYEELYNKTPQLOLDAFEVEKQALLNEHGAQEQ 300
Qy 619 LNKIRDSYAKLLGHONLKQIKHVYVKLDENSOLKSEVSKRLCQAKKK 667
Db 301 LNKIRDSYAKLLGHONLKQIKHVYVKLDENSOLKSEVSKRLCQAKKK 349

RESULT 6

US-08-402-217A-3
Sequence 3, Application US/08402217A
Patent No. 5587301
GENERAL INFORMATION:
APPLICANT: Hawkins, Phillip R.
Wilde, Craig G.

APPLICANT: Seilhamer, Jeffrey J.
TITLE OF INVENTION: HYALURONAN RECEPTOR EXPRESSED IN HUMAN
TITLE OF INVENTION: UMBILICAL VEIN ENDOTHELIAL CELLS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/402,217A
FILING DATE: 10-MAR-1995
CLASSIFICATION: 435.
ATTORNEY/AGENT INFORMATION:
NAME: Luther, Barbara J.
REGISTRATION NUMBER: 33954
REFERENCE/DOCKET NUMBER: PF-00280US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-852-0195
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 477 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
IMMEDIATE SOURCE:
LIBRARY: mouse
CLONE: GI 53979
US-08-402-217A-3

Query Match 38.2%; Score 1374.5; DB 1; Length 477;
Best Local Similarity 60.8%; Pred. No. 2.7e-70;
Matches 291; Conservative 33; Mismatches 68; Indels 87; Gaps 2;

Qy 319 MONLKQKFTLEQOEHEKLOQKELQIDSLLOQKEKLSLSLHQKLCFSQOEVMKEKNLFEE 378
Db 1 MONLKQKFTLEQOEHEKLOQKELQIDSLLOQKEKLSLSLHQKLCFSQOEVMKEKNLFEE 60
Qy 379 LKQTLDELKLOQKEQAEERLVKQLEEEAKSRAEBLKLEELKLGKGAELKSAHAHTQA 438
Db 61 LKQTLDELKLOQKEQAEERLVKQLEEEAKSRAEBLKLEELKLGKGAELKSAHAHTQA 120
Qy 439 TLLLOEKYDSMVOSLEDVTAQESYKALTASIEDLKLENSLQEKAAKAGNAEDVQHQ 498
Db 121 TLLLEKYDSMVOSLEDVTAQESYKALTASIEDLKLENSLQEKVAKAGNAEDVQHQ 180
Qy 458 ILATSSNQYVYMLDLQTKSALKETEITVTSFLQKITDLOLQKQOEEDFRKQLED 558
Db 181 ILATSSNQYVYMLDLQTKSALKETEITVTSFLQKITDLOLQKQOEEDFRKQLED 240
Qy 559 EGRKAENTTAETETINKWRLLYEELYNKTPQLOLDAFEVEKQALLNEHGAQEQ 618
Db 241 EGRKAENTTAETETINKWRLLYEELYNKTPQLOLDAFEVEKQALLNEHGAQEQ 300
Qy 619 LNKIRDSYAKLLGHONLKQIKHVYVKLDENSOLKSEVSKRLCQAKKK 667
Db 301 LNKIRDSYAKLLGHONLKQIKHVYVKLDENSOLKSEVSKRLCQAKKK 349

Db 421 EVSKLSOLVRKQNELRLOGELDKALGIRHFDPSKAFCHASKENF---TPLKEGPNPC 476

RESULT 7

US-08-700-178-3

Sequence 3, Application US/08700178

Patent No. 5783669

Patent No. 5783669 5700912

GENERAL INFORMATION:

APPLICANT: Hawkins, Phillip R.

APPLICANT: Wilde, Craig G.

APPLICANT: Seilhamer, Jeffrey J.

TITLE OF INVENTION: HYALURONAN RECEPTOR EXPRESSED IN HUMAN

TITLE OF INVENTION: UMBILICAL VEIN ENDOTHELIAL CELLS

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA: US/08700,178

APPLICATION NUMBER: US/08700,178

FILING DATE: August 20, 1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/402,217

FILING DATE: March 10, 1995

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0028-1 DIV

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 477 amino acids

TYPE: amino acid

STRANDEDNESS: single

MOLECULE TYPE: protein

IMMEDIATE SOURCE:

LIBRARY: mouse

CLONE: GI 53979

US-08-700-178-3

Query Match 38.2%; Score 1374.5; DB 1; Length 477;

Best Local Similarity 60.8%; Pred. No. 2.7e-70;

Matches 291; Conservative 33; Mismatches 68; Indels 87; Gaps 2;

Qy 319 MQLKQKFIHQEHEKLOQKELQIDSLLOQEKELSSSLHOKLCSFQEMVKEKNLFEE 378

Db 1 MQLTERLALEREVEKLOQKELQSLQSLQKELSLARLOQLCSFQEMTSEKNVFKEE 60

Qy 379 LKOTLDELKLOQKEQAEARLVKOLEEAKSRAEKLKLEKLGKAELEKSSAAHTQA 438

Db 61 LKLAELDAVQKQEEQERLVKQLEERKSTAEQLTRDNLLEKEVELEKHIAHAQA 120

Qy 439 TILLQEKYDSMVQSLDYVT----- 457

Db 121 ILIAQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESEFQEKYNDTAQSLRDV 180

Qy 458 -----AQFESYKALTASEIDL 474

Db 181 TAQLESEKYNDAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESYKSSSTLKEIDL 240

Qy 475 KLENSLOEKAAGKNAEDVQHOILATESSSNOEYVRMLDLQTKSALKETKEITVSF 534

Db 241 KLENTLQEKVAMAESVEDVQOQILTAESTNOEYARVWQDLQNRSTLKEEIKEITVSF 300

Qy 535 LOKITDLQNLKQOEDFRKQLEDEEGRAEKENTTAELTEINKWRLLYEELYNKTKPF 594

Db 301 LEKITDLKNLQROQDEFRKQLEEKGRKTAENYMTLMTBINKWRLLYEELYEKTKPF 360

Qy 595 QLOLDAFEVEKQALLNEHCARQEQNLKRDYSYAKLLGHONLQKOKKHVVKLKDENSOLKS 654

Db 361 QOQLDAFEAEKQALLNEHGATQEQNLKRDYSYAKLLGHONLQKOKKHVVKLKDENSOLKS 420

Qy 655 EVSKLRCOLAKKKQSETKLOEELNKVLGKTHKFDPSKAFHESKENFALKTPLKEGNTNC 713

Db 421 EVSKLSOLVRKQNELRLOGELDKALGIRHFDPSKAFCHASKENF---TPLKEGPNPC 476

RESULT 8

US-08-995-654-3

Sequence 3, Application US/08995654

Patent No. 6025138

GENERAL INFORMATION:

APPLICANT: Hawkins, Phillip R.

APPLICANT: Wilde, Craig

APPLICANT: Seilhamer, Jeffrey

TITLE OF INVENTION: HYALURONAN RECEPTOR EXPRESSED IN HUMAN

TITLE OF INVENTION: UMBILICAL VEIN ENDOTHELIAL CELLS

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Dr.

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows

SOFTWARE: FastSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/995,654

FILING DATE: December 22, 1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/700,178

FILING DATE: August 20, 1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/402,217

FILING DATE: March 10, 1995

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0028-2 DIV

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 477 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-995-654-3

Query Match 38.2%; Score 1374.5; DB 3; Length 477;

Best Local Similarity 60.8%; Pred. No. 2.7e-70;

Matches 291; Conservative 33; Mismatches 68; Indels 87; Gaps 2;

Qy 319 MQLKQKFIHQEHEKLOQKELQIDSLLOQEKELSSSLHOKLCSFQEMVKEKNLFEE 378

Db 1 MQLTERLALEREVEKLOQKELQSLQSLQKELSLARLOQLCSFQEMTSEKNVFKEE 60

Db	61	TSSFLEKIDTLKQLRQQDEDFRKQLEEKGTAEKENVMTELTMEINKWRLLYBELYEK	120
Qy	591	TKPFLQLDAFEVEKQALLNEHGAAQEQLNKIRDSYAKLLGHONLKQKTKHVVKLKDNS	650
Db	121	TKPFPQQLDAFEAKQALLNEHGATQEQLNKIRDSYAQLLGHONLKQKTKHVVKLKDNS	180
Qy	651	QLK 653	
Db	181	QLK 183	

```

RESULT 11
US-09-572-191-2
; Sequence 2, Application US/09572191
; Patent No. 6355466
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6355466el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1017
; CURRENT APPLICATION NUMBER: US/09/572,191
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1388
; TYPE: PRT
; ORGANISM: Human
US-09-572-191-2

```

```

Query Match      11.5%; -Score 413; DB 4; Length 1388;
Best Local Similarity 21.0%; Pred. No. 9.4e-16;
Matches 171; Conservative 157; Mismatches 282; Indels 206; Gaps 25;

QY 37 SFKRSQRFKO-----QKESQNINLND-----K 58
| : : ||: | |||:: |
Db 598 SKQBYEFKELTRKRQLESELSQKANUNLENLENATKACRQEVQSUNKHTAETLK 657
QY 59 DTTLPASARKVKSSEKES-----QKNDKDLKILEK 90
| : : | | | | |
Db 658 IITPTKRAYQLHSRPVKLPSPMGSGSLYTONSSILDNDILNFPVPENNEQAFALSE 717
QY 91 EIRVLQERGAQDRRIOD-----LTELEKME-----ARLNALREKYSLS 131
| : | : | : | : | : | : | : | : | : | : |
Db 718 ELRTVQPMQALQAKDEEHEKNKLKLOHQVHDKLEHSHSTQMOELFSSERIDWTQKQOEELS 777
QY 132 ANNATLEKQILIELTRTNELKLSKFSE-----NGNQKRLILSLMLKLRNKRRTKMRGMM 186
| | | | | | | | | | | : | : | : | : | : | : |
Db 778 QLN-VLEKQLOETQTKNDFLKVSEVHDURVVLHSADKELSSVKLEYSSFKTNQKEFNKL- 835
QY 187 AKQEGEMKLVQTVQRSLE-----ESQG-----KLAQLEGKLVSTIEKPKID 226
| : | : | : | : | : | : |
Db 836 -SERMHVQLQDLNRLNENEKLESKACIQSDSYDNLQELMKFEIDQLSRNLQNFKENET 894
QY 227 EKSETEKLEYIEBISCASDOVEYKIDIAQLENLEKNDIEILSKQSLDEENTVILSKQ 286
| : | : | : | : | : | : | : | : | : | : |
Db 895 LKSDLNINLMELLE-----AEKERNKLSL-QFEEDKENSKEILKVLVAVRQ-----EKQ 943
QY 287 VEDLNVAQCQ-----LLEKEKE-DHVNENRREHNENLNAEMONLKQKFILEQOE 333
| | | : | : | : | : | : | : | : | : | : |
Db 944 KE--TAKCEQMAMKVQKLESRLATEKVISSLEKSRSDRKKVADLMNQIQEL-----R 995
QY 334 EKLOQKEQLDLSLQOKEBELSSSLHOKLCSFQEMVK----- 370
| : | : | : | : | : | : |
Db 996 SSVCEKTETIDTLKQELKDINCYNALSALVDREESRVLLKQOEVDILDKETLRLILSED 1055
QY 371 -ENKLFPEELKQTLDELDKLQOKEEQEARLVKOLLEBBAKSRAEBKLLEKCLKQKEAELE 429
| : : | : | : | : | : | : | : | : | : | : |
Db 1056 IERDMLCEDLAHATEQOLNMLTAEASKKSHGSLQSAQOELTKKEALTOELHOKLNOKKKEVE 1115
| : : | : | : | : | : | : | : | : | : | : |

```

```

Qy 430 KSSAHTQATLLLOEKYDSWVQSLED--VTAOFESYKA-----LTASETEDUKLENSLSQE 483
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1116 QKNEYNFKMRQLEHYMDSAADPPQSPKTPPHFOHLAKLLETQBOETEDGRASKTSLKH 1175
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 484 KAAGKAGNAEDVQHOILATSESSNOEYVRMLLDQ---TKSALKETEIKEITVVSFLQKITD 540
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1176 LVTKLNEDEVKNAELIRMKQEURLRENMENLRLESQQLIEKNWLLQQLDDIK---RQKENS 1232
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 541 LQN-----QLKQOQEDPRKO-----LEDEGRKAKEKNTTAELTEBIN 578
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1233 DONHPDNOQLKNEQESIKERLAKSKIIVEMLKMRADLEEVSQALYNKEMECLRMWTDVE 1292
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 579 KRWLLTYEELYNKTKPQOLDAFEVEKQALLNEHGAQEOQLNKIRDSYAKLLIGHONLKQK 638
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1293 RTOTLESKAFQEKEQURSKLEEMYERERTSOEEMLRKQVECLAENGKLVGHONLHQK 1352
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 639 IKHVVLKLDNSOLKSEVSKLRQC---LAKKROSET 671
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1353 IQYVRLKKNVRLAEETEKLRANVFLAKEKRSES 1388
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
US-09-723-262-2
; Sequence 2, Application US/09723262
; Patent No. 6379912
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6379912el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1017
; CURRENT APPLICATION NUMBER: US/09/723,262
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: US 09/572,191
; PRIOR FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows version 4.0
; SEQ ID NO 2
; LENGTH: 1388
; TYPE: PRT
; ORGANISM: Human
US-09-723-262-2

```

Query Match	11.5%;	Score	413;	DB	4;	Length	1388;
Best Local Similarity	21.0%;	Pred.	No. 9.4e-16;				
Matches	171;	Conservative	157;	Mismatches	282;	Indels	206;
Gaps							25;
Qy	37	SFKSQRFKQ-----QKESQNUNVD-----	K	58			
		: :	:				
Db	598	SKQVEYFKELTRKQLESELSQAKNLNLENLEATKACKRQEVSQLNKIHAETLK	657				
Qy	59	DTTLPASARKVKSSEKES-----QKNDKDLKILEK	90				
		: :	:				
Db	658	IITPTKAYQILSPRPVKLSPMGFSGLYQNSIILDNDILNEPVPMEHQEAFaise	717				
Qy	91	EIRVLLQERGADRRIQD-----LETELEKME-----ARLNAALREKTSLS	131				
		: :	: :				
Db	718	ELRTVQSQMSALQAKLDEEHKNKLQOHVDKLEHSTQMQLFSSERIDMTKQOEBLLS	777				
Qy	132	ANNATLEKQLLELRTTNELLSKFSE-----NGNKNLRILSLLEMLRKNKRETKMRGMM	186				
		: :	: :				
Db	778	QLN-VLEKQLOQETQKDNDFLAKSEVHDLRVVLHSADKELSSVKFSYFKTNOKEFNKL-	835				
Qy	187	AKQEGEMKMLQVTRQSLE-----ESOG-----KTAQLEGLKLVSTEKEKID	226				
		: : :	: :				
Db	836	-SERHHVQLQDLNRLUNENEKLLSEKACLDQSYONLQEIIMKFEIDQLSRNLQNFKKENET	894				
Qy	227	EKSETEKLEVEIBISCASDOVEKYKLDIAQLEENLEKNDKELSLKSLSEENIVILSKQ	286				
		: : :	: :				
Db	895	LKSDLNINLMELLE-----AEERNKKLSL-QFEEDKENSKEILKVLVAVRQ-----EKQ	943				
Qy	287	VEDLNVRACQ-----LLEKEKE-DHVNENRHNENLNAEMONLQKQFLEQOEH	333				

RESULT 15

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:09:43 ; Search time 36.3569 Seconds
(without alignments)
4108.822 Million cell updates/sec

Title: US-09-685-010-47
Perfect score: 3598
Sequence: 1 MSFPKAPLKRFPDPSGCAPS.....LKEGNTNCRAPMECQESWK 725

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organella.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2670.5	74.2	713	11 Q9WUF7	Q9wuf7 rattus norv
2	2623	72.9	836	11 Q8VDR2	Q8vdr2 mus musculu
3	1366	38.0	476	11 Q92026	Q92026 mus sp. hya
4	996	27.7	240	6 Q8SPR5	Q8spr5 bos taurus
5	633	17.6	182	11 Q9CX55	Q9cx55 mus musculu
6	444	12.3	1790	3 Q07380	Q07380 saccharomyc
7	442	12.3	2139	5 Q07569	Q07569 entamoeba h
8	436.5	12.1	1690	5 Q44929	Q44929 drosophila
9	429.5	11.9	1411	4 Q15075	Q15075 homo sapien
10	425.5	11.8	1410	4 Q14221	Q14221 homo sapien
11	418	11.6	1944	13 Q9DGM5	Q9dgm5 gallus gall
12	416	11.6	1738	5 Q76329	Q76329 dictyosteli
13	414	11.5	1940	6 Q9BE41	Q9be41 bos taurus
14	413	11.5	1598	11 Q922D2	Q922d2 mus musculu
15	412	11.5	1388	4 Q9NS87	Q9ns87 homo sapien
16	411.5	11.4	746	4 Q9H6N6	Q9h6n6 homo sapien

17	411	11.4	1941	13 Q9DGM4	Q9dgm4 gallus gall
18	409	11.4	1368	13 Q9PTD7	Q9ptd7 xenopus lae
19	409	11.4	2007	13 Q02015	Q02015 gallus gall
20	408.5	11.4	1583	4 Q15045	Q15045 homo sapien
21	408	11.3	1939	6 Q9TV63	Q9tv63 sus scrofa
22	408	11.3	2017	5 Q94992	Q94992 drosophila
23	408	11.3	2057	5 Q94987	Q94987 drosophila
24	407.5	11.3	1937	13 Q9IBD4	Q9ibd4 gallus gall
25	407.5	11.3	1960	11 Q8VDD5	Q8vdd5 mus musculu
26	407	11.3	2056	11 Q9W088	Q9w088 drosophila
27	406.5	11.3	1690	5 Q9VJES	Q9vje5 drosophila
28	406	11.3	1937	6 Q9TV62	Q9tv62 sus scrofa
29	406	11.3	1938	6 Q9BE40	Q9be40 bos taurus
30	405.5	11.3	1964	13 Q93522	Q93522 xenopus lae
31	403	11.2	1939	6 Q9TV61	Q9tv61 sus scrofa
32	402	11.2	1463	5 Q9GY20	Q9gyz0 strongyloce
33	402	11.2	2138	5 Q9AZE3	Q9aze3 amoeba prot
34	400.5	11.1	1177	16 Q8RCY8	Q8rcy8 thermoanaer
35	400.5	11.1	1625	13 Q90Y12	Q90y12 gallus gall
36	400.5	11.1	1939	13 Q9PTV2	Q9pty2 gallus gall
37	400	11.1	1999	11 Q63731	Q63731 rattus norv
38	399.5	11.1	1538	13 Q98940	Q98940 gallus gall
39	399	11.1	1931	13 Q91973	Q91973 coturnix co
40	398.5	11.1	804	4 Q8TC31	Q8tc31 homo sapien
41	398	11.1	1391	11 Q922J3	Q922j3 mus musculu
42	397.5	11.0	1945	6 Q97757	Q97757 felis silve
43	396.5	11.0	2115	4 Q14980	Q14980 homo sapien
44	396	11.0	1909	5 Q25893	Q25893 plasmodium
45	395.5	11.0	2101	4 Q14981	Q14981 homo sapien

ALIGNMENTS

RESULT 1

ID	Q9WUF7	PRELIMINARY:	PRT;	713 AA.
AC	Q9WUF7;			
DT	01-NOV-1999 (Tremblrel. 12, Created)			
DT	01-NOV-1999 (Tremblrel. 12, Last sequence update)			
DT	01-JUN-2001 (Tremblrel. 17, Last annotation update)			
DE	Hyaluronan receptor RHAMM.			
GN	RHAMM.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=SPRAGUE DAWLEY;			
RA	Li X., Lynn B., Nagy J.I., Cattini P.A.;			
RT	"RHAMM cDNA from rat brain."			
RL	Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=PHOCHROMOCYTOMA;			
RA	Lynn B.D., Li X., Cattini P.A., Nagy J.I.;			
RT	"Sequence, protein expression and erk association of the hyaladherin RHAMM in PC12 cells."			
DR	EMBL; AF133037; AAD24473.1; -			
DR	EMBL; AF336825; AAK21904.1; -			
KW	Receptor.			
SQ	SEQUENCE 713 AA: 82395 MW; 2FE310D5759C6CB3 CRC64;			

Query Match	74.2%;	Score 2670.5;	DB 11;	Length 713;
Best Local Similarity	75.3%;	Pred. No. 9.8e-88;		
Matches 537;	Conservative 71;	Mismatches 104;	Indels 1;	Gaps 1;

Oy	1	MSFPKAPLKRFPDPSGCAPSPGAYDKVTLEVLKGPVSFOKSORFKQOKESKONLNVDKDT	60
Db	1	MSFPKAPLKRFPDPSGCAPSPGAYDKVTSESTKGPVSFOKSORFKNQRESQONLNIDKDT	60

```
QY 61 TLPASARKYKSSKESKQNDKDLKLEKEIRVLVQLQERGAQDRRIQDLETELEKMEARL 120
DB 61 TLLASAKAKTLYSKKESKQNDKDKVRLKEKEIRVLVQLQERGTQDKRIQDMSELEANTEANL 120
QY 121 NAALREKTSLSANNATLEKQLTETLTNTNELLKSKSFSENGNOKNRLTSLSELMKLNKRET 180
DB 121 NAAVTEKPSLSANNASLEKRLTETLTNELLKSKSFSENGNOKNRLTSLSELMKLNKRET 180
QY 181 KWRGMAKQEGHEMQLQVTORESLEESOGKIAOLEGKLVSTIEKEKIDSEKTEKLLIYIEE 240
DB 181 KWRGMAKQEGHEMQLQVTORESLEESOGKIAOLEGKLVSTIEKEKIDSEKTEKLLIYIEE 240
QY 241 ISCASDQVEKYKLDIAQLEENLKEKNDKDEILSLKQSLSEENIVILSKOVEDLNKVCQLEKE 300
DB 241 ISCASDQVEKYKLDIAQLEENLKEKNDKDEILSLKQSLSEENIVILSKOVEDLNKVCQLEKE 300
QY 301 KEDHVRNREHEHNENLNAEMNLKQKILFEOQHEKLOQKELQIDSLQOKEKELSSSLHOK 360
DB 301 KEDHVRNREHEHNENLNAEMNLKQKILFEOQHEKLOQKELQIDSLQOKEKELSSSLHOK 360
QY 361 LCSFQEMVYKERNLFEELKQTLDELQKLOQKEQAERLVKQLEEBKASRAEBELKLEEK 420
DB 361 LCSFQEMVYKERNLFEELKQTLDELQKLOQKEQAERLVKQLEEBKASRAEBELKLEEK 420
QY 421 LKGEAELEKSSAAHTQATLL----- 441
DB 421 LKGEAELEKSSAAHTQATLL----- 441
QY 481 LQKAAKAGKNAEDVOHILATNESSNOEYVRLMLDLQTKSALKETEIKETVTFLOKIND 540
DB 481 LQKAAKAGKNAEDVOHILATNESSNOEYVRLMLDLQTKSALKETEIKETVTFLOKIND 540
QY 541 LQNLQKQEDFRKQLEDEBGRKAENKTAELTEINKRWLLYEELYNKTKPPFQQLDA 600
DB 541 LQNLQKQEDFRKQLEDEBGRKAENKTAELTEINKRWLLYEELYNKTKPPFQQLDA 600
QY 601 FEVEKQALLNEHGAQEQNLKIRDSYAKLLGHQNLKQKIKHVVKLKDENSQKSLR 660
DB 601 FEVEKQALLNEHGAQEQNLKIRDSYAKLLGHQNLKQKIKHVVKLKDENSQKSLR 660
QY 661 COLAKKQKQETKLOEELNKVLGKHKHDPKAFHESKENFALTKPLKEGNTNC 713
DB 661 COLAKKQKQETKLOEELNKVLGKHKHDPKAFHESKENFALTKPLKEGNTNC 713
QY 660 SQAKRQKQELRLQGLDRLGELDKALGIRHFDPSKAFCHESKENVTLTKPLKEGNPC 712
DB 660 SQAKRQKQELRLQGLDRLGELDKALGIRHFDPSKAFCHESKENVTLTKPLKEGNPC 712
RESULT 2
Q8VDR2 ID Q8VDR2 PRELIMINARY; PRT; 836 AA.
AC Q8VDR2;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE Similar to hyaluronan mediated motility receptor (RHAMM).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC021427; AAH21427.1; -.
KW Receptor.
SQ SEQUENCE 836 AA; 96670 MW; B9BDDA22BECA652 CRC64;
Query Match 72.9%; Score 2623; DB 11; Length 836;
Best Local Similarity 65.1%; Pred. No. 5.5e-86;
Matches 546; Conservative 66; Mismatches 97; Indels 130; Gaps 3;
QY 1 MSFPKAPLARNDPSCAPSPGAYDVKTLEVLKGPVSFKOSRFKQKESKONLWVDKDT 60
DB 1 MSFPKAPLARNDPSCAPSPGAYDVKTSEATKGPVSFKOSRFKQKESKONLWVDKDT 60
```

```
QY 61 TLPASARKYKSSKESKQNDKDLKLEKEIRVLVQLQERGAQDRRIQDLETELEKMEARL 120
DB 61 TLLASAKAKTLYSKKESKQNDKDKVRLKEKEIRVLVQLQERGTQDKRIQDMSELEANTEANL 120
QY 121 NAALREKTSLSANNATLEKQLTETLTNTNELLKSKSFSENGNOKNRLTSLSELMKLNKRET 180
DB 121 NAAVTEKPSLSANNASLEKRLTETLTNELLKSKSFSENGNOKNRLTSLSELMKLNKRET 180
QY 181 KWRGMAKQEGHEMQLQVTORESLEESOGKIAOLEGKLVSTIEKEKIDSEKTEKLLIYIEE 240
DB 181 KWRGMAKQEGHEMQLQVTORESLEESOGKIAOLEGKLVSTIEKEKIDSEKTEKLLIYIEE 240
QY 241 ISCASDQVEKYKLDIAQLEENLKEKNDKDEILSLKQSLSEENIVILSKOVEDLNKVCQLEKE 300
DB 241 ISCASDQVEKYKLDIAQLEENLKEKNDKDEILSLKQSLSEENIVILSKOVEDLNKVCQLEKE 300
QY 301 KEDHVRNREHEHNENLNAEMNLKQKILFEOQHEKLOQKELQIDSLQOKEKELSSSLHOK 360
DB 301 KEDHVRNREHEHNENLNAEMNLKQKILFEOQHEKLOQKELQIDSLQOKEKELSSSLHOK 360
QY 361 LCSFQEMVYKERNLFEELKQTLDELQKLOQKEQAERLVKQLEEBKASRAEBELKLEEK 420
DB 361 LCSFQEMVYKERNLFEELKQTLDELQKLOQKEQAERLVKQLEEBKASRAEBELKLEEK 420
QY 421 LKGEAELEKSSAAHTQATLL----- 441
DB 421 LKGEAELEKSSAAHTQATLL----- 441
QY 481 LQKAAKAGKNAEDVOHILATNESSNOEYVRLMLDLQTKSALKETEIKETVTFLOKIND 540
DB 481 LQKAAKAGKNAEDVOHILATNESSNOEYVRLMLDLQTKSALKETEIKETVTFLOKIND 540
QY 541 LQNLQKQEDFRKQLEDEBGRKAENKTAELTEINKRWLLYEELYNKTKPPFQQLDA 600
DB 541 LQNLQKQEDFRKQLEDEBGRKAENKTAELTEINKRWLLYEELYNKTKPPFQQLDA 600
QY 601 FEVEKQALLNEHGAQEQNLKIRDSYAKLLGHQNLKQKIKHVVKLKDENSQKSLR 660
DB 601 FEVEKQALLNEHGAQEQNLKIRDSYAKLLGHQNLKQKIKHVVKLKDENSQKSLR 660
QY 661 COLAKKQKQETKLOEELNKVLGKHKHDPKAFHESKENFALTKPLKEGNTNC 713
DB 661 COLAKKQKQETKLOEELNKVLGKHKHDPKAFHESKENFALTKPLKEGNTNC 713
QY 660 SQAKRQKQELRLQGLDRLGELDKALGIRHFDPSKAFCHESKENVTLTKPLKEGNPC 712
DB 660 SQAKRQKQELRLQGLDRLGELDKALGIRHFDPSKAFCHESKENVTLTKPLKEGNPC 712
RESULT 3
Q92026 ID Q92026 PRELIMINARY; PRT; 476 AA.
AC Q92026;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE Hyaluronan receptor.
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-92299690; PubMed-1376732;
RA Hardwick C., Hoare K., Owens R., Hohn H.P., Hook M., Moore D.,
RA Cripps V., Austen L., Nance D.M., Turley E.A.;
RT "Molecular cloning of a novel hyaluronan receptor that mediates tumor
cell motility."
RL J. Cell Biol. 117:1343-1350(1992).
```


RP	SEQUENCE FROM N.A.
RX	MEDLINE=95286647; PubMed=7768953;
RA	Mu F.T., Callaghan J.M., Steele-Mortimer O., Stenmark H., Parton R.G.,
RA	Campbell P.L., McCluskey J., Yeo J.P., Tock E.P., Toh B.H.;
RT	"EAL, an early endosome-associated protein. EAL is a conserved
RT	alpha-helical peripheral membrane protein flanked by cysteine
RT	'fingers' and contains a calmodulin-binding IQ motif.;"
RL	J. Biol. Chem. 270:13503-13511(1995).
DR	EMBL; L40157; AAA79121.1; .
DR	InterPro; IPR000822; Znf_C2H2.
DR	InterPro; IPR00306; Znf_FYVE.
DR	Pfam; PF01363; FYVE; 1.
DR	SMART; SMO0355; Znf_C2H2; 1.
DR	PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
DR	PROSITE; PS00157; ZINC_FINGER_C2H2_2; 1.
KW	DNA-binding; Zinc-finger.
SQ	SEQUENCE 1410 AA; 162496 MW; DF0F9464D70A8AED CRC64;
	Query Match 11.8%; Score 425.5; DB 4; Length 1410;
	Best Local Similarity 25.6%; Pred. No. 6.1e-08;
	Matches 185; Conservative 141; Mismatches 264; Indels 133; Gaps
Qy	30 EVLKGPFVSQKRSQRKQKESKONLVNDKTTLPASARKVKSSESKSOKNDDKLKILE 89
Db	I : :
Db	275 ELSGPQEVAVVQELQKLKSSVNLTQKNQTILTNLLKKKEQDYTKLEEKHNEE--SVSK 332
Qy	90 KEIRVLLOERGAQDRRIODLETELEKMEARLNAAALREKTSLSANNATLEKQLIELTTRNE 149
Db	: :
Db	333 KNIQATLHQK-----DLDCQQLOSRLSAS----ETSLH----RIHVSEKGEATQ 375
Qy	150 LKKSFSENGNGNLRILSLMLKRNKRRTK-MRCGMAKQE--GMEMKLQVTFORSLEES 206
Db	: :
Db	376 KLKEELSE--VETKYQHAKAEFKQEQQREEKEQHGLQLOEINQHLKLETTEROLGEA 433
Qy	207 QGK-----IALEGKLVSIIE---KEKIDSEKTEKLEYEE-----240
Db	: :
Db	434 HGRLKEQRLSSEKLMDEKQVADLQLKSLRLEEQLKEKVTNSTELHQHLDKTKQHQEQ 493
Qy	241 ---ISCASDOVEKYKLDIAOLEENLEKENDELISLK---QSLEENIVILSKOVEDLNVC 294
Db	: :
Db	494 QALQOSTTAKLEAGANDLEVLRIGEEKDOTNLQALLQKSKENTSLLEKEREDLYAKI 553
Qy	295 OLLEKEKEHDVNRNHNENLNAEMONLKOKFTLEQOEHEKIQKQELQIDLSQLQEKELS 354
Db	: :
Db	554 QAGEGETA-VLNQLEKNHTLQDEVTLTEN-VKNQSESHKQAQENLH-DOVQEQKAHLR 610
Qy	355 SSLHQKLCFSQEMYKKNLFEEELKQTLDELDKLQOKEEQEAERLVKYLEEAKSRAE-- 412
Db	: :
Db	611 AAOQDVLS--LETSVNELNSQLNESKEKVSOLD--IQIKAKTELILLSAABAATAQADLQ 666
Qy	413 -ELKILLEKLGCKEAELEKSSAAHTQATILLQEKYDSVMQSLVEDVTAQEFYSKALTASEI 471
Db	: :
Db	667 NHLDTAQNALQKHQELNKITTLQDPVTAQLQDKQEHCSHLEKYEKYLSSL-EQKT 735
Qy	472 EDL-----KLENSSLQEKAAKAGNAEDVOHO-----ILATESNO-----EYVRML 513
Db	: :
Db	726 EELEGQIKKLEADSLEVKASKE-QALQDLQOQQROLNTDLELRATELSKOLEMEKEIVSST 784
Qy	514 -LDLQTSALKETEIKEITYSFLOKITDLQNQLKQOEBDFRKOLEDGEGRKAEKNTTAE 572
Db	: :
Db	785 RLDAQKS-----EALESIQQKLT-----KQEEE--KQILKQDFETLSQET---- 823
Qy	573 LTEETNKWRLLYEELYNYTKPFLQDLDAFEVEKQALLNEHGAAEQOLNKIRDYS----- 626
Db	: :
Db	824 -----KIOHEELNNRIQTTVTTELQKVMEKEALMTSELSTVKDKLSVSDSLKNSKE 875
Qy	627 -----AKLLGHONLKOKIKHHVVKLDENSOLKSEVSKLRCOLAKKQSETKLQEEL 677
Db	: :
Db	876 FEKENQKGAAILDLEKTCCKELHQLOVQMONT-LK-EQELKLSLEKEREASHQULLEL 933
Qy	678 NKV 680
	:

SQ	SEQUENCE 1411 AA; 162465 MW; C3B17777FE34B6BD CRC64;
	Query Match 11.9%; Score 429.5; DB 4; Length 1411;
	Best Local Similarity 25.2%; Pred. No. 4.4e-08;
	Matches 183; Conservative 146; Mismatches 257; Indels 141; Gaps 31;
Qy	30 EVLKGPFVSQKRSQRKQKESKONLVNDKTTLPASARKVKSSESKSOKNDDKLKILE 85
Db	I : :
Db	275 ELARGPQEVAVVQELQKLK----SSVNLTQKNQTILTNLLKKKEQDYTKLEEKHNEE-- 328
Qy	86 KILKEIRVLLOERGAQDRRIODLETELEKMEARLNAAALREKTSLSANNATLEKQLIELT 145
Db	: :
Db	329 SVSKKNTOATLHQK-----DLDCQQLOSRLSAS----ETSLH----RIHVSELSKG 371
Qy	146 RTNELLSKFSENGNGNLRILSLMLKRNKRRTK-MRGMAKQE--GMEMKLQVTFORS 202
Db	: :
Db	372 EATOKLKEELSE--VETKYQHAKAEFKQEQQREEKEQHGLQLOEINQHLKLETTERQ 429
Qy	203 LEESQGK-----IALEGKLVSIIE---KEKIDSEKTEKLEYEE-----240
Db	: :
Db	430 LGAEHRGLKEQRLSSEKLMDEKQVADLQLKSLRLEEQLKEKVTNSTELHQHLDKTKQ 489
Qy	241 -----ISCASDOVEKYKLDIAOLEENLEKENDELISLK---QSLEENIVILSKOVEDL 290
Db	: :
Db	490 HQEQALQOSTTAKLEAGANDLEVLRIGEEKDOTNLQALLQKSKENTSLLEKEREDL 549
Qy	291 NVKCOLLEKEKEHDVNRNHNENLNAEMONLKOKFTLEQOEHEKIQKQELQIDLSQLQOE 350
Db	: :
Db	550 YAKTOAGEGETA-VLNQLEKNHTLQDEVTLTEN-VKNQSESHKQAQENLH-DOVQEQKAHLR 610
Qy	351 KELSSSLHQKLCFSQEMYKKNLFEEELKQTLDELQKQOEQAERLVKQLEBEAKSR 410
Db	: :
Db	606 KAHLRAAQDRVLSL-ETSVNELNSQLNESKEKVSOLD--IQIKAKTELILLSAABAATAQ 662
Qy	411 AE--ELKILLEKLGCKEAELEKSSAAHTQATILLQEKYDSVMQSLVEDVTAQEFYSKALT 467
Db	: :
Db	663 ADLQNHLDTAQNALQKHQELNKITTLQDPVTAQLQDKQEHCSHLEKYEKYLSSL- 721
Qy	468 ASETEDL-----KLENSSLQEKAAKAGNAEDVOHO-----ILATESNO-----EY 509
Db	: :
Db	722 EQKTEELGGQIKKLEADSLEVKASKE-QALQDLQOQQROLNTDLELRATELSKOLEMEKEI 780


```
Db 789 QYKQOOLS-SNSNIDQQ-----LQSTIISELSEKELNDKSLIEKEKOLQOQEFDOQL 843
QY 103 DRRIQ-DLETELEKEARLNAALREKTSLSANNATLEKOLIELTTRTNELLKSKFSENGNQ 161
Db 844 NEKNQKHQDQLELEKQLOQOQYDQLNQNETNQNLNQNOLNK-----ENLNE 896
QY 162 KNLRIILSLMLKRNKRETKMRGMMAKQEGMEMKLQVTRORSLEESQGKIAQLEGLVSTE 221
Db 897 K-----EQELLQNLQNOQIEKIQDQOEFQKNSINIELVNEKNEKLIQLOQDYDQLK 951
QY 222 KE--KIDKSETEKILEYIEEISCASQVQEKYKLDIAQLEENLKKNDEILSKQSLSEEN 279
Db 952 QOQRNSNDKEND-LIE-----KENOLKSTQNELNQLIEKNESDHKEQOQLKQOSIEND 1003
QY 280 IYVLSKQVEDLVNKKCOLLEK-----EKEDHVNRRNREHNENLNAEMONLKQKFILEQ 330
Db 1004 LIEKENQIQQL--QSLNEQROQQNQSLSEKQOQLNQLIEKQ--FDQKEQOQLKQOSIEND 1060
QY 331 --QEHEKIQOKELOIDSLQOQEKELSSLHQKLSFQOEMVYKKNLFEELKOTLDLQK 388
Db 1061 LFEKENQIQQLQSLNEQROQQNQSLSEKQOQLNQLIEK--NESQKQOQLKQOSIEND- 1117
QY 389 LQOKEEQARLVKQLEEEAKSRAE-----ELKLEELKKGKEAELEKSSAAHTQATLL 442
Db 1118 LIEKENQIQQLQSLNEQROQQNQSLSEKQOQLNQLIEK--NESQKQOQLKQOSIEND 1177
QY 443 QEK---YDSM-----VQSLVEDYTAQFESYKALTASEIEDLKLNSLQEK 484
Db 1178 QDKQIEFDQLQTLTFNQFNKNDKDSQFIQLODDQKQOQLSIQ-----QDLNOLKQOENQEKQ 1233
QY 485 AAKAGNAEDVQHLATYESSNQEVYRMLDLQTRKALKETEKITVSLFKITDLO-- 542
Db 1234 LSEKDEKLIQSIFQNEQKEKQSLSEKDEKLIQSIOQNLQNLNDENQEKVQFSEKDEKLQSI 1293
QY 543 ----NOLKQOEDFRKQLEDEGRKAENKTAETAEINKWRLLYELYNKTKPFQLOL 598
Db 1294 QOQLNOLKQOENQEKQOL-----SEKDEKLIQSIOQDLNQ--LNDQIKKNEKXK-- 1340
QY 599 DAFEVEKQALLNEHGAQPOLNKRDSYAKLIGHONLAKOKIKHVYKLDKENSOLKSE--V 656
Db 1341 ---EKEBQLLKQDFNDQOQSLKQLEKSEKENQIQOQLKQOENIOLNQQOQSNEII 1397
QY 657 SKLRQLAKKQOSE-----TKLQBEL-----NKVLGI-----KHFD 687
Db 1398 QOLKQOQLKQOQEQOQENNEKEIERLQIEQLKQOQOIEQSELSNEIKIQTTFQOQED 1457
QY 688 PSKAFHSHESKENFALK 703
Db 1458 --QLSHNRKSKDQLHLQ 1471

RESULT 13
Q9BE41 PRELIMINARY; PRT; 1940 AA.
AC Q9BE41;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Myosin heavy chain 2a.
GN MYHC-2A.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_taxid=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HOLSTEIN; TISSUE=SKELETAL MUSCLE;
RA Chikuni K., Muroya S., Nakajima I.;
RT "Sequencing of the bovine myosin heavy chain isoforms.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB059398; BAB40920.1; -.
```

```
DR HSSP; P13538; 2MYS.
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR001609; myosin_head.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR PRODOM; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS00096; IQ; 1.
SQ SEQUENCE 1940 AA; 223318 MW; EE2642ELE29FDDC2 CRC64;

Query Match 11.5%; Score 414; DB 6; Length 1940;
Best Local Similarity 23.4%; Pred. No. 2.le-07;
Matches 205; Conservative 135; Mismatches 261; Indels 276; Gaps 36;

QY 51 KQNLNVKDDTTLPASARKVKSESKEKQNDKDLKILEKEIRV-----LLO 97
Db 774 KQYKVLNASAIPEG-----QYIDSKKASEKLLASIDIDHTQYKFGHTKVFVKAGLLGLLE 779
QY 98 ERGAQDRRIQDLETELEKEARLNAALREKTSLSANNATLEKQLIELTTRTNELLKSKFSE 157
Db 780 E--MRDEKLAQLMT---RTQARCRCGLAR-----VEYQKMWVERRESIEFC- 818
QY 158 NGNQKNLR-----ILSLELMKLNKRETKMRGMMAKQOEMK--LQVTRORSLEESQGKIA 211
Db 819 --IQYNIRAFNVKHPWNKLFRIKPLKLSAETEKEMATKKEEPQKTDELAKSEAKRK 876
QY 212 QLEGKLVIEKEIDEKSETKLEVEIEISCASQOQVKYL-----DIAQLEENLKE 264
Db 877 ELEEKWVTLLEKNDLQLOVQSEAGLADAEBCDQIKTKIQLBAKKEVTERADEEE 936
QY 265 KNDETLSLKQLEENIVILSKQVEDLVNKKQLEKEKEDHVNRRNREHNENLNAEMONL-- 322
Db 937 INAEITAKRKLEDECSSELKDDIDDLTLETLAKVEKEK--HATENKV--KNLTEEMAGLDE 992
QY 323 -----KQKFILE-----QOHEKQLQ-----QKELQIDSL--LQOKEKELSS 356
Db 993 TIAKTKKKAQEAHQOQLDDQAEEDKVNTLTAKTKLEQOVDLDEGSLQOQEKKLKRD 1052
QY 357 LHO-----KLCSFOEMV---KEKNLFEELKQTLDELKQLOK--EBOA----- 396
Db 1053 LERAKRKLEGLDKLA--QESIMDINEKQQLDEKLKKEFEISNLQSKIEDQALGIQLQ 1110
QY 397 -----ERLVKQLEEEAKSRAE-----ELKLLPEKL-----K 422
Db 1111 KKIKELQARIEELEEEIEAERASRAKAEKQSDLSRELEEEISERLEEAGGATSAQIEMNK 1170
QY 423 KEAE-----LEKSSAAHTQATLLQEKY--DSM-----VQSLVEDYTAQFESYKALTA 468
Db 1171 KREAEFQKMRDLEATLQHEATAALKRKHADSVAEELGEQIDNLQRYKQLEKEKSEM 1230
QY 469 SEIEDKLKLENSLOEKAACAKNAEDVOHQLATYESSNQEVYRMLDLQTRKALKETEI- 527
Db 1231 MEIDDLASNVETISKAKGNLEKMCRTLEDQVNLKSKKEEQORLINDLTQGRLOQESG 1290
QY 528 -----KEITVS-----FLQKITDLOLQKQO-----BED---PRK 554
Db 1291 EFSRQLEDEKALVSQLSRGKQAFQOIEELKQLEEEIKAKNALAHQLSARHDCDLLRE 1350
QY 555 QLEDEEGRKAKENNTAELTEINKWRLLY-----EELYNKTKPFQLOLDA----- 600
Db 1351 QYEEEOESKAEQLORALSKANTEVAQWRTKYETDAITQRTTEELEEAKKKLAQRLQAEHVE 1410
QY 601 -----FEVEKQALLNE-----HGAQPOLNKRIRDSYAKLLG-----HQ 633
Db 1411 AVNKAASLETKORLQNEVEDLMDVETNACALDKORNFQKILAEWKQKYEETHA 1470
QY 634 NL-----KQIKHVYKLDKENSOLKSEVSKRLCOLAK----- 665
```



```
Db 1471 ELEAAQKEARSLGTFLFKMKNAYEESLDQLETLKRENKNLQOEISDLTEQIAEGGKRMHE 1530
Qy 666 ----KKQ-----SETKLOEELNKLVLGK 684
Db 1531 LEKIKKQVEQKSEIQAALEAEASLEHEEGKILRIQ 1567

RESULT 14
Q922D2
ID Q922D2 PRELIMINARY; PRT; 1598 AA.
AC Q922D2;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Similar to myosin, heavy polypeptide 2, skeletal muscle,
DE adult..
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC008538; AAH08538.1; -.
DR InterPro; IPR000048; IQ.region.
DR InterPro; IPR001609; myosin_head.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR PROSITE; PS00096; IQ; 1.
SQ SEQUENCE 1598 AA; 183083 MW; DE896B34346A479 CRC64;

Query Match 11.5%; Score 413; DB 11; Length 1598;
Best Local Similarity 24.8%; Pred. No. 1.9e-07;
Matches 184; Conservative 114; Mismatches 205; Indels 240; Gaps 30;

Qy 172 MKLRNRETKRMGMKQEGMEMK--LQVTRSLSEESQCKTAQKGLVSEKEKIDKS 229
Db 837 MKLFFKIRPLKLSAETEKEMATMKEEFQTKDOLAKSEAKRKELEKVVSLKERNDLQL 896
Qy 230 ETEKLEVEIEISCASDOVEYKL-----DIAOLEENLKEKNDEILSLKQSLLENIVI 282
Db 897 QVQAAEGLADAERCDQLIKTKIOLEAKIKEVTERAEEDEEINAEALTAKKRKLEDCGE 956
Qy 283 LSKOVEDLNVRKCOLLEKEKEDHVNRENHNENLNAMQNL-----KQKFILE----- 329
Db 957 LKKDIDDLLELTIAKVEKEK--HATENKV--KNLTTEAGLDETIAKLTKEKKAQEAHQ 1012
Qy 330 -----QOEHEKIQ-----QKELQIDSL--LQOEKELSSLSHQ-----KLCFQ 365
Db 1013 TLDDLQAEEDKQNTLTAKIKLEQQVDDLEGSLEQKLRWDLERAKRKLEGLDLKA--Q 1070
Qy 366 EEWV---KEKNLFEELKOTDELDKLOOK--EEQA-----ERLVKOLE 404
Db 1071 ESIMDIENEKQOLDRKAKKEFEMSNLOSKEIDEQAIGTQKQKIKELQARTEEELEETE 1130
Qy 405 EEAKSRAE-----ELKILLEKL-----KGKAE-----LEKSSA 433
Db 1131 AERASRAEAKQORSOLSRLEEEIEISRELEAGATSAQIEMNKKRAEAFQKMRDRLEEAATL 1190
Qy 434 AHTQATLLIQEY--DSM-----VQSLDVTQAQFYSKALTASEIEDLKLENSLQEKAA 486
Db 1191 QHEATAATLRKHAHSVLAELGEQIDNLRVKQKLEKESEKMEIDDL-----ASNVEIVS 1246
Qy 487 KAGKNAE----DVQHQILATESNQQEYVRMLDLQTKSALKETEI-----KEITVS 533
Db 1247 KAKGNLEKMCRTLDQVSELKSEEEQOORLINDLTSQSRQLQTSEGFSTRQJDEKEALVS 1306
Qy 534 -----FLOKITDLOLKOQOED-----FRKQLEDESGRAEKEN 568
Db 1307 QLSRGKQAFQOIEELKQLEEEVAKNALAHALOSSRHDCCDLLREQYEEESKAELQR 1366
Qy 569 TTAELTEEINKWRLLY-----EELYNKTKPFOLOLDA-----FEVEKQ 606
Db 1367 ALSKANSEVAQWRKYETDAIQRTTEELBEAKKLAQRLQAEEHVEAVNAKCALEKTKQ 1426
Qy 607 ALLNE-----HGAQEQOLNKIRDSYAKLLG-----HQNL----- 635
Db 1427 RLQNEVEDLMLDVERTNAACAALDKQRNDFKILAEWKQYETHAELEASKEARSIGT 1486
Qy 636 -----KQIKHWVKLKDENSOLSKSEVSKLRQLAK-----KKQ----- 668
Db 1487 ELFKMKNAYEESLDQLETLKRENKNLQOEISDLTEQIAEGGKRIHELEKIKKQVEQKCE 1546
Qy 669 -----SETKLOEELNKLVLGK 684
Db 1547 LQAALAEAEASLEHEEGKILRIQ 1569

RESULT 15
Q9NS87
ID Q9NS87 PRELIMINARY; PRT; 1388 AA.
AC Q9NS87;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Kinesin-like protein 2.
GN HKLP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20435852; Pubmed=10878014;
RA Sueishi M., Takagi M., Yoneda Y.;
RT "The forkhead-associated domain of Ki-67 Antigen Interacts with the
RT Novel Kinesin-like Protein Hk1p2.";
RL J. Biol. Chem. 275:28888-28892(2000).
DR EMBL; AB035898; BAB03309.1; -.
DR HSP; p17119; 3KAR.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW ATP-binding; Coiled coil; Microtubules; Motor protein.
SQ SEQUENCE 1388 AA; 160160 MW; E127EB4B9991CA83A CRC64;

Query Match 11.5%; Score 412; DB 4; Length 1388;
Best Local Similarity 21.0%; Pred. No. 1.8e-07;
Matches 171; Conservative 157; Mismatches 282; Indels 206; Gaps 25;

Qy 37 SFQKQSQRFQ-----QKESKNLVND-----K 58
Db 598 SKQEEFEKELTRKROLESELSQKLANLENLEATKACKRQEVSQLNKHIAETUK 657
Qy 59 DTTLPASARKVKSSSEKES-----QKNDKDLKILK 90
Db 658 ITTTTKAYQLHSRVPVKLSPWGSFGSLYTONSSITLDNILEPVPPEMQAFAISE 717
Qy 91 EIRVLQERGAODRIQD-----LETELEKME-----ARUNAAERKTSUS 131
Db 718 ELURTVOEQMSALQALDDEEHKNLKLQHVQKLEHHSTOMQELFSSERIDWTQKQEEILS 777
Qy 132 ANNATLEQOLIELTITNELLKSKFSE-----NGNQKNLILSLMLKLNKRETKRMGM 186
Db 778 QLN-VLEKQLQETQTKDNDFLKSEVHDLRVVLHSADEKELSSVKLEYSSSFKTNOKEFENKL- 835
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:06:53 ; Search time 10.6932 Seconds
(without alignments)
2812.095 Million cell updates/sec

Title: US-09-685-010-47

Perfect score: 3598

Sequence: 1 MSFPKAPLKRFPNDPGCAPS.....LKEGNTNCRAPMECQESWK 725

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3568.5	99.2	724	1	HMMR_HUMAN
2	2623	72.9	794	1	HMMR_MOUSE
3	1768.5	49.2	498	1	HMMR_RAT
4	444	12.3	1790	1	USO1_YEAST
5	422	11.7	2230	1	GOG4_HUMAN
6	419	11.6	1976	1	MYHA_RAT
7	417	11.6	1875	1	MLP1_YEAST
8	414	11.5	1976	1	MYHA_HUMAN
9	414	11.5	2116	1	MYSD_DICDI
10	409.5	11.4	1939	1	MYH4_HUMAN
11	409	11.4	1959	1	MYH9_CHICK
12	408	11.3	2017	1	MYSN_DROME
13	407.5	11.3	1433	1	REST_CHICK
14	407	11.3	1941	1	MYH2_HUMAN
15	405	11.3	1679	1	YIO9_YEAST
16	404.5	11.2	1960	1	MYH9_HUMAN
17	403.5	11.2	1938	1	MYSS_CHICK
18	403	11.2	1938	1	MYH4_RABIT
19	403	11.2	1940	1	MYH3_RAT
20	402.5	11.2	1972	1	MYHB_RABIT
21	402	11.2	1978	1	MYHB_CHICK
22	400.5	11.1	1940	1	MYH3_CHICK
23	398.5	11.1	1972	1	MYHB_HUMAN
24	398	11.1	1102	1	MYSC_CHICK
25	398	11.1	1940	1	MYH3_HUMAN
26	398	11.1	1972	1	MYHB_MOUSE
27	398	11.1	1976	1	MYHA_BOVIN
28	396	11.0	1937	1	MYH8_HUMAN
29	395.5	11.0	1427	1	REST_HUMAN
30	395	11.0	1961	1	MYH9_RAT
31	394.5	11.0	1938	1	MYH9_HUMAN
32	393	10.9	976	1	SCP1_HUMAN
33	393	10.9	1939	1	MYH1_HUMAN

RESULT 1

ID	HMMR_HUMAN	STANDARD;	PRT;	724 AA.
AC	075330; Q92767;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Hyaluronan mediated motility receptor (intracellular hyaluronic acid binding protein) (Receptor for hyaluronan-mediated motility) (CD168 antigen)			
DE	HMMR OR IHABP OR RHAMM.			
GN	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND CHARACTERIZATION.			
RC	TISSUE=Breast carcinoma;			
RX	MEDLINE=98264864; PubMed=9601098;			
RA	Assmann V., Marshall J.F., Fieber C., Hofmann M., Hart I.R.;			
RT	"The human hyaluronan receptor RHAMM is expressed as an intracellular protein in breast cancer cells.";			
RL	J. Cell Sci. 111:1685-1694(1998).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Breast;			
RX	MEDLINE=97045829; PubMed=8890751;			
RA	Wang C., Entwistle J., Hou G., Li Q., Turley E.A.;			
RT	"The characterization of a human RHAMM cDNA: conservation of the hyaluronan-binding domains.";			
RL	Gene 174:299-305(1996).			
CC	-1- FUNCTION: INVOLVED IN CELL MOTILITY. WHEN HYALURONAN BINDS TO HMMR, THE PHOSPHORYLATION OF A NUMBER OF PROTEINS, INCLUDING THE FOCAL ADHESION KINASE OCCURS. MAY ALSO BE INVOLVED IN CELLULAR TRANSFORMATION AND METASTASIS FORMATION, AND IN REGULATING EXTRACELLULAR-REGULATED KINASE (ERK) ACTIVITY.			
CC	-1- SUBUNIT: SUBUNIT OF THE HARC COMPLEX.			
CC	-1- SUBCELLULAR LOCATION: CELL SURFACE AND CYTOPLASMIC (BY SIMILARITY).			
CC	-1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A (SHOWN HERE) AND B; ARE PRODUCED BY ALTERNATIVE SPLICING.			
CC	-1- TISSUE SPECIFICITY: EXPRESSED IN BREAST CANCER CELL LINES AND IN NORMAL BREAST TISSUE.			
CC	-1- DATABASE: NAME=PROW; NOTE=PROW 2:76-84(2001); WWW="http://www.ncbi.nlm.nih.gov/prow/guide/802868666.g.htm".			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AF032862; AAC32548.1; -.			
DR	EMBL; U29343; AAC52049.1; -.			

P54697 dictyosteli
Q62209 mus musculu
P12883 homo sapien
P13533 homo sapien
P13539 mesocricetu
Q90339 cyprinusu ca
Q02566 mus musculu
Q03410 rattus norv
O67124 atafex aeo
Q10411 schizosacch
P49454 homo sapien
Q99996 h a-kinase

ALIGNMENTS

CC TRANSFORMATION AND METASTASIS FORMATION, AND IN REGULATING
CC EXTRACELLULAR-REGULATED KINASE (ERK) ACTIVITY (BY SIMILARITY).
CC -!- SUBUNIT: SUBUNIT OF THE HARC COMPLEX.
CC -!- SUBCELLULAR LOCATION: CELL SURFACE AND CYTOPLASMIC (BY
CC SIMILARITY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; U87983; AAB47997.1; -.
KW Hyaluronic acid; Repeat; Glycoprotein.
FT DOMAIN 420 430 HYALURONIC ACID-BINDING (POTENTIAL).
FT DOMAIN 442 451 HYALURONIC ACID-BINDING (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 302 302 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 483 483 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 498 AA; 57858 MW; 58037C79BD5C2A70 CRC64;

Query Match 49.2%; Score 1768.5; DB 1; Length 498;
Best Local Similarity 73.1%; Pred. No. 2.2e-51;
Matches 356; Conservative 51; Mismatches 79; Indels 1; Gaps 1;

QY 227 EKSETEKLEYIEEISCASDQVEKYKLDIAOLENKEKNDRIILSKQSLEENIVILSKQ 286
DB 12 EKSETEKLEYIEEISCASDQVEKYKLDIAOLENKEKNDRIILSKQSLEENIVILSKQ 70

QY 287 VEDLVNKKCOLLEKEKEDHVNREHNENLNAEMQNLKQKFILEQEPHEKLOQELQIDSL 346
DB 71 IEDLVNKKCOLLEARDLVSKDRERAESLSAEMQVLTETKLLERQEKLOQNELQSLSL 130

QY 347 LQOEKELSSHLQKCSFOEMVMEKKNLFEELKOTLDLQKQKEQAEQLVQLQLEEE 406
DB 131 LQOEKELSAHQQLQCSFOEEMTSERNVPKEQLKALDELDAVQOQKEQSEKLVQLEEE 190

QY 407 AKSRAPELKLEKLGKAELEKSSAAHTQATLLLOEKYDSWQSLQEDVTAQFESYKAL 466
DB 191 TSTAEQLRLDOLLEKEIELEKTAHAQAQVTAQEKYSTAQTLRDVTAQLESYKSS 250

QY 467 TASEIEDLKLENSLQEKAAKAGNAEDVOHILATESNQYVRLMLDLQTKSALKETE 526
DB 251 TLKEIEDLKLENLTQEKVMAEKVEDVQOQILTAESTNQYAKVQDQLQNSSTLKEAE 310

QY 527 IKEITVSLQKTDLQNLQKQEEEDFRKOLEDEEGKAEKENTTAELTEENKWRLLYEE 586
DB 311 IKEITVSLQKTDLQNLQKQEEEDFRKOLEDEEGKAEKENTTAELTEENKWRLLYEE 370

QY 587 LYNKTKPFQQLDAPFEVKEQALLNHEGAQOLNKTROSYAKLGHONLKQIKHVVKLK 646
DB 371 LYDKTKPFQQLDAPFEVKEQALLNHEGAQOLNKTROSYAKLGHONLKQIKHVVKLK 430

QY 647 DENSLQKSEVSLRQKAKKQSETKLOBELNKLVLGIRHFDPSKAFHHESKENFALKTPL 706
DB 431 DENSLQKSEVSLRQKAKKQSETKLOBELNKLVLGIRHFDPSKAFHHESKENFALKTPL 490

QY 707 KEGNTNC 713
DB 491 KEGNPNC 497

RESULT 4
USOL_YEAST
ID USOL_YEAST STANDARD: PRT; 1790 AA.
AC P25386;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Intracellular protein transport protein USOL.
GN USOL OR INT1 OR YDL058W.

OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=X2180-1A;
RX MEDLINE=91185402; PubMed=2010462;
RA Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,
RA Yamasaki M.;
RT "A cytoskeleton-related gene, usol, is required for intracellular
RT protein transport in Saccharomyces cerevisiae.";
RT J. Cell Biol. 113:245-260(1991).
RN [2]
RP SEQUENCE OF 782-1790 FROM N.A.
RA Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,
RA Kendrick K.E.;
RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-8 FROM N.A.
RA Bai Y., Symington L.S.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI
CC COMPLEX.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR
CC MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE
CC ER AND THE GOLGI COMPLEX.
CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED
CC OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL
CC COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.
CC -!- SIMILARITY: BELONGS TO THE VDP/USOL/YBL047C FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X54378; CAA38253.1; -.
DR EMBL; L03188; AAB00143.1; -.
DR EMBL; U53668; AAB66659.1; -.
DR PIR; A38455; A38455.
DR SGD; S0002216; USOL.
DR InterPro; IPR002017; Spectrin.
KW Transport; Protein transport; Golgi stack; Cytoskeleton; Coiled coil.
FT DOMAIN 1 724 GLOBULAR HEAD.
FT DOMAIN 725 1790 COILED COIL (POTENTIAL).
FT DOMAIN 465 487 CHARGED (HYPER-HYDROPHILIC).
FT DOMAIN 991 1790 DISPENSABLE FOR THE PROTEIN FUNCTION.
FT DOMAIN 1172 1786 ASP/GLU-RICH (ACIDIC).
FT CONFLICT 847 847 G -> E (IN REF. 2).
FT CONFLICT 924 924 E -> K (IN REF. 2).
FT CONFLICT 1253 1253 V -> I (IN REF. 2).
FT CONFLICT 1319 1319 I -> V (IN REF. 2).
FT CONFLICT 1461 1461 N -> S (IN REF. 2).
FT CONFLICT 1581 1581 G -> S (IN REF. 2).
FT CONFLICT 1600 1600 I -> V (IN REF. 2).
FT CONFLICT 1661 1661 R -> S (IN REF. 2).
FT CONFLICT 1772 1772 D -> DEEDDEE (IN REF. 2).
SQ SEQUENCE 1790 AA; 206424 MW; 6CE2B216E9FD4818 CRC64;

Query Match 12.3%; Score 444; DB 1; Length 1790;
Best Local Similarity 23.3%; Pred. No. 4.9e-08;
Matches 187; Conservative 166; Mismatches 277; Indels 174; Gaps 32;

QY 44 FKQKESQNLNVDKDITLPASARKVSKSEKES-----QKNDKDKI--LEKEIRV 94
DB 946 YKRSFQSHDNLVAKLTKELKLSLANNYKDMQAEKNESLIKAVEESKNESLQNLQNKIDS 1005

QY 95 LLQ-----ERGAQDRRIQ-----DLTELEKMEARLNA-----LREK--T 128
GN USOL OR INT1 OR YDL058W.

```
Db 1006 MSQKENFQIBRSGISIEKNIQKTKTISDLEQTKKEIIISKDSKDEYESQISLKEKLET 1065
Qy 129 SLSANATLEKOLJELTRTNELLSKFSSENGNOKNRLISLEMLKRNKRETKRMGMMAK 188
Db 1066 ATTANDENVK-ISELTKYTRLEAEALAAAYKLNK------ELETKLETSEKALKEV 1115
Qy 189 QEG-----MEMKQVQTSLSQKIAQLEGLKYSIEKEKID-----EKSETEKILE 236
Db 1116 KENEHLKEEKIQLEKEKATETKQ-QLNSLRANLESLEKEHEDLAQALKYEEQIANKERQ 1174
Qy 237 YIEISCASDOVEYKLDIAQLENLEKND-----EILSLKQSLBENIVILSKQVEDLNV 292
Db 1175 YNEISQINDBITSTQ-----QENESIKKKNDLEGEYKAMKSTSEQSNLKKSEIDALNL 1230
Qy 293 KCOLLEKEKEDHVNRNREHNENNAEQNLKQKFILOQEQEHEKLOQKELOIDSLLOQKE 352
Db 1231 QIKELKKKNETNEASLESISKVSSEVVK-----ELQDCNFKKEV-----SELEDKLK 1282
Qy 353 LSSSLHOKLCSFQEMVKEKNLFE---BELKQTLDELKLOQKEQAE-----RLVKOLEE 405
Db 1283 ASEDKNSKYLEQSEKIEBELDAKTELKIQLEKITNLSKAREKSESELSRLKKTSE 1342
Qy 406 EAKSRAELKLEKLEKKEAELEKSSAHTQATLLOQEKYDMSVQSLDVTAFESYKA 465
Db 1343 ERKNAEQLEKLEKWEIQKNOAFKERKLLNEGSTITQOEYSEKINTLEDELIRLQNE 1402
Qy 466 L-----TASEIDKLLENSSLOEKAAGKAGNAED----- 494
Db 1403 LKAKEIDNTRELEKVSLSNDELLEEQNTIKSLQDELLSYKDKITRNDKLLSIEDNKK 1462
Qy 495 -----VQHILATSSNOEYVRLMLDLOTKSALKETEI---KEITVSFLQKITDLQNLK 546
Db 1463 RDLSEKQLRAAQESKAKVEGLKLEESSEKAELEKSEMMKLESTIESNETLK 1522
Qy 547 QOEDDFKQLED-BEGRKAKE-----NTTAEITELINKWRLIYEELYNTKPFQOLD 599
Db 1523 SSMETIRSKDKLSQSKSAEEDTKNQHEKSDLSIRNESEKIDIELKSK-----LRIE 1577
Qy 600 A-----FEVEKQALNEHGAQEQNLKIRDSYAKLLGHQNL-LKQKIKHVYK-LKDENSQ 651
Db 1578 AKSGSELETVKQLNN---AQEKI-----RINAENIVLKSLEDIERELKDKQAE 1625
Qy 652 LK-----SEVSKLRCLAKKQSETKLQBELKNVGIKHPDFSKAFHESKENFA 701
Db 1626 IKSQOEKELLTSRLKELEQLDSTQQAQKSEEE---RRAEVKRFQVEKSLD----- 1677
Qy 702 LKPLKEGNTNCRYAPMEQCSWK 725
Db 1678 -KAMLETK-----YNDLVNKEQAWK 1697

RESULT 5
GOG4_HUMAN STANDARD; PRT; 2230 AA.
AC Q13439; Q14436; Q13270; Q13654;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Golgi autoantigen, golgin subfamily A 4 (Trans-Golgi p230) (256 kDa
DE golgin) (Golgin-245) (72.1 protein).
GN GOLGA4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=96215236; PubMed=8626529;
RA Erlich R., Gleeson P.A., Campbell P., Dietzsch E., Toh B.-H.;
RT "Molecular characterization of trans-Golgi p230: a human peripheral
RT membrane protein encoded by a gene on chromosome 6p12-22 contains
RT extensive coiled-coil alpha-helical domains and a granin motif.";
RL J. Biol. Chem. 271:8328-8337(1996).
```

```
[2]
SEQUENCE FROM N.A.
Seelig H.P.;
Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE OF 131-2230 FROM N.A.
TISSUE=Placenta;
MEDLINE=96125112; PubMed=8537393;
Fritzler M.J., Lung C.-C., Hamel J.C., Griffith K.J., Chan E.K.L.;
"Molecular characterization of golgin-245, a novel Golgi complex
protein containing a granin signature.";
J. Biol. Chem. 270:31262-31268(1995).
[4]
SEQUENCE OF 524-672 FROM N.A.
TISSUE=Gastric fundus;
Balague C.;
Thesis (1994), Instituto municipal de investigacion medica, Spain.
-1- FUNCTION: MAY PLAY A ROLE IN VESICULAR TRANSPORT FROM THE TRANS-
GOLGI.
-1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE CYTOPLASMIC FACE OF THE
GOLGI MEMBRANE.
-1- ALTERNATIVE PRODUCTS: At least 4 isoforms: 1 (shown here), 2, 3
and 4; are produced by alternative splicing.
-1- DISEASE: ANTIGEN IN THE AUTOIMMUNE DISEASE SJOEGREN'S SYNDROME AND
IN HEPATITIS B.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; U41740; AAC50434.1; -
EMBL; X82834; CAAS8041.1; -
EMBL; U31906; AAC51791.1; -
EMBL; X76942; CAAS4261.1; -
Gene; HGNC:4427; GOLGA4.
MIM: 602509; -
InterPro; IPR000237; GRIP_domain.
Pfam; PF01465; GRIP; 1.
Golgi stack; Antigen; Coiled
DOMAIN 133 237
DOMAIN 276 1011
DOMAIN 1033 1214
DOMAIN 1259 2152
VARSPIC 2154 2185
VARSPIC 2186 2230
VARSPIC 2103 2109
VARSPIC 2222 2230
FT CONFLICT 188 188
FT CONFLICT 220 220
FT CONFLICT 276 276
FT CONFLICT 584 584
FT CONFLICT 628 628
FT CONFLICT 630 630
FT CONFLICT 682 682
SQ SEQUENCE 2230 AA; 261139 MW; 3BB733DB1EA86134 CRC64;
Query Match 11.7%; Score 422; DB 1; Length 2230;
Best Local Similarity 22.8%; Pred. No. 3.1e-07;
Matches 216; Conservative 151; Mismatches 275; Indels 306; Gaps 39;
Qy 26 VKTLEVLKGPVSFQKSRFKQKQ-----KESQNLNVDKDTTLTPASARKVKSSEKESQK 80
Db 278 VKTLETLQ-----QVRKQENLLKRCETIQSHKQCITLLTSEKALQELDERLQE 329
Qy 81 ND--KDLKILKE-----IRVLQERG---AQDRR-----TQDLETELEKWEA 118
Db 330 LEKIKDLHMAEKTKLITQLRDAKNLIEQLEQDKGMVIAETKQRMHETLEMKEEIAQLRS 389
```



```

QY 193 EMKLVQTSLEESOGKTAQLEGLKLVSTIEKEIDEKSETKLEKLEYIEISCASQDQVEKYK 252
Db 1314 ESQLODTQELLQETROKLNLSRRQLEEEK-----NSLQEQEEREARKNQV 1366
QY 253 LDIAQLEENLEKKEDEILSLKQSLLE-----NIVILSKQVEDLVNVCQLEKEKEDHV 305
Db 1367 LALQSLQADTKKKYDDDDGTTGEEFAKKKLLKQVEALSQRLEEKVLAYDKLETK----- 1422
QY 306 NRNEHNENNAEONLKKQFLEQOEHEKLOQKELOLQDSLQOQKELSSLSHOK----- 360
Db 1423 NRLQQLDLDLVLDH-----QRQIVSNLEKKQKQKFDQLLAEKIGISARYAERDRAE 1475
QY 361 -----LCSFQEMVKEKNLFEELKQTLDELQKLOQKEBEQAERLVKQLEEPK 408
Db 1476 AEAREKETKALSARALBEALEAKEEFERQNKQLRADMEDLMSSKDDVGNVHELEKSKR 1535
QY 409 SRAEELKLEKLGKEAELEKSSNAHTQATLILLOKQYDSMVQSLEEDVTAQFESYKALTA 468
Db 1536 ALBQOQVEEMRTQLELEDELEQADAKLRLEVNMQAKQAQFERDLQTRDEQNEEKKRLLL 1595
QY 469 SEIEDL--KLENSLSQEKAAAGK-----NAEDVQHQILATESNOEYVRMLDLQTKSA 521
Db 1596 KQVRELEAELEDEKQRALAVASKKWEIDLKLEAQLEAANKARDEVIKQLKQAOQWK 1655
QY 522 LKTEIKETIVTS-----FLQ-----KITDLQNLKQOE-----EDFRKOLED 558
Db 1656 DYQRELEEARSRDEIFAQSKSEKSKLLEAEILQLEELASSERRARRHAEQERDELAD 1715
QY 559 E-----EGRKA---EK---ENTTAELTEINKWRLLYELLYNKKYFPQOLDADFEVQKA 607
Db 1716 ETANSASGKSALLDEKRLLEARIQAOLEEELQESNMELLNDRFKTTLOYVDTLNTE--- 1772
QY 608 LLNEHGAQEQNLKTRDSYAKLLGHON--LKOKTKHV-----VKLKDNSOLSKSEVSKLR 660
Db 1773 LAEERSAAQKS-----DNARQOLERQNKELAKQLEEGAVKSKFKATISALEAGIGOLE 1827
QY 661 CQL---AKKKOSETKQLEELNKVL-----GIKFHPDSPKAFHESKENF--ALKTPLK 707
Db 1828 EQLEQEAERAAANKLVRRTEKKLEIFPMQVEDERRHADQYKEQMEKANARMKQLKROLE 1887
QY 708 EGNTCNYRA 716
Db 1888 EAEETATRA 1896

RESULT 7
MLP1_YEAST
ID MLP1_YEAST STANDARD; PRT; 1875 AA.
AC Q02455;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin-like protein MLP1.
GN MLP1 OR YKR095W OR YKR415.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RX MEDLINE=93247549; PubMed=8483450;
RA Koelling R., Nguyen T., Chen E.Y., Botstein D.;
RT "A new yeast gene with a myosin-like heptad repeat structure.";
RL Mol. Gen. Genet. 237:359-369(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94205265; PubMed=8154186;
RA Bou G., Esteban P.F., Baladron V., Gonzalez G.A., Cantalejo J.G.,
RA Renacha M., Jimenez A., del Rey F., Ballesta J.P.G., Revuelta J.L.;
RT "The complete sequence of a 15,820 bp segment of Saccharomyces
RT cerevisiae chromosome XI contains the UB12 and MLP1 genes and three
RT new open reading frames.";
```

```

RL Yeast 9:1349-1354(1993).
CC -!- FUNCTION: MYOSIN-LIKE PROTEIN THAT IS PROBABLY INVOLVED IN DNA
CC REPAIR.
CC -!- SIMILARITY: SOME, TO THE TPR ONCOGENE.
CC -!- CAUTION: REF.2 MISQUOTES THE GENE NAME AS "MPL1".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; L01992; AAA34783.1; -
DR EMBL; X73541; CAA51948.1; -
DR EMBL; Z28320; CAA82174.1; -
DR PIR; S38173; S38173.
DR SGD; S0001803; MLP1.
KW Coiled coil; DNA repair.
FT DOMAIN 531 1678
FT DOMAIN 1834 1866
FT CONFLICT 301 301 R -> A (IN REF. 1).
SQ SEQUENCE 1875 AA; 218455 MW; 683A0D34C9066867 CRC64;

Query Match 11.6%; Score 417; DB 1; Length 1875;
Best Local Similarity 24.1%; Pred. No. 3.8e-07;
Matches 190; Conservative 145; Mismatches 29; Indels 162; Gaps 30;

QY 29 LEVLKGPVSFQSRFKQKESQNLNVD---KDTTLPASARKVKSSEKES--QKND 82
Db 816 LSELAKETS-QKDHHIKQLEED--NNSNIEWQNKIEALKKDYESTVTSQKTDIEKIQ 873
QY 83 KDLKLEKEI---RVLLQERGAQDRRIQD--LETELEKWEALNAA-----LREKT 128
Db 874 YKVSLEKEIEEDKIRLHTYNYMDETINDDSLKLEKSKINLTDAYSQIKYKDIYETT 933
QY 129 SISANNATLEKQILIELTRTNELLSKFSQNGQ--KNL-----RILSLELMKLAN 176
Db 934 SOS-----LQOTNSKLDSEDFDTNQIKNLTDEKYSLEDKISLLEKEQMFNLNN 981
QY 177 KRETKRMGMMAQEGEMKLVQTSLEESQKIAQLEGLKLVSTIEKEIDEKS----- 229
Db 982 ELDLQKKGMEKEKADFKKRISILQNNKEVEAVKSEYESKLSKIQND--LDQOTIYANTAQ 1040
QY 230 -----ETEKLEYIEISCASQDQVEKYK-----LDIAQLEENLEKKEDEILSLKQSL 277
Db 1041 NNYEQELQKHADVSKTISELREQLHTYKQGVKTLNLSRDLQENALKENESKSSQKESLL 1100
QY 278 ENIVILSKQVEDLVNVCQLEKEKEDHVNRNHNENLNA-----EMQNLKOK 325
Db 1101 EQLDLSNRIEDLSQNLKLLYDQIQIYTAADKEVNNSNTPGGLNNLITLRERDILDTK 1160
QY 326 FTLEQOEHEKQKELQIDSLQO-----EKELSSLSHQKLCSPQEE-----MYK 370
Db 1161 VYVAERDAKMLRQKISLMDVELQDARTKLDNSRVEKENHSSIIQQHDDIMEKLNQLNLLR 1220
QY 371 EKNL--FEEL-----KOTLDELKQKEEQEAERLVKQLEBEAKSRAEELKLEKL-- 421
Db 1221 ESNITLRLENNNKKEKLQSELQDLKQNVAPIESELTAALKYSMOEKQELKLAKEEVH 1280
QY 422 ---KGEAELEKSSAAHTQATLILLOKQYDSMVQSLE-----DYTAQFESYKALT 467
Db 1281 RWKKRSQDITLQK-----HEQLSSSDYKDESEIENLEKEENKERQGAEEKFNRLRQA 1336
QY 468 ASEIEDLKLENSLSQEK-----AKAGKNAEDVQHQILATESNOEYVRML 513
Db 1337 QERLTKLSQDSLTQEVNLSRDAKNVLENSLEARNIEELQNAKVAQGNQLEAIRKL 1396
QY 514 L-DLQTKSALKTEIKETIVTSFLQKITDLQNLKQOEEDFRKOLEDDEGRKAEEKNTTAAE 572
Db 1397 QBDAAKASRELQAKLEESTTSVESTINGLNEETITTLKEIEKQROIQOQQLQATSANEQND 1456
```


Db 1773 LAERSAAQKS-----DNARQQLERQNKELKAKIOLEGAVKSKFKATISALEAKIGOLE 1827
 Qy 661 CQL---AKKQSEKTKLOBELNKVL-----GKHKFDPKSKAPHESKENF-ALKTPLK 707
 Db 1828 EQLSEAKERAANKLVRRTEKKLEIFMQVEDERRHADQYKEQMEKANARKQLKROLE 1887
 Qy 708 EGNTNVCYRA 716
 Db 1888 EAEEEAATRA 1896

RESULT 9
 MYS2_DICDI
 ID MYS2_DICDI STANDARD; PRT; 2116 AA.
 AC P08799;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin II heavy chain, non muscle.
 GN MHCA.
 OS Dictyostelium discoideum (slime mold).
 OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87092266; PubMed=3540939;
 RA Warrick H.M., de Lozanne A., Leinwand L.A., Spudich J.A.;
 RT "Conserved protein domains in a myosin heavy chain gene from
 Dictyostelium discoideum.";
 RL Dictyostelium discoideum.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:9433-9437(1986).
 RN [2]
 RP PHOSPHORYLATION SITES, AND MUTAGENESIS.
 RC STRAIN=AX2;
 RX MEDLINE=90353583; PubMed=2387408;
 RA Lueck-Vielmeier D., Schleicher M., Grabatin B., Wippler J.,
 RA Gerisch G.;
 RT "Replacement of threonine residues by serine and alanine in a
 phosphorylatable heavy chain fragment of Dictyostelium myosin II.";
 RL FEBS Lett. 269:239-243(1990).
 RN [3]
 RP PHOSPHORYLATION SITES.
 RX MEDLINE=88112226; PubMed=2828113;
 RA Magle G., Noegel A., Scheel J., Gerisch G.;
 RT "Phosphorylation of threonine residues on cloned fragments of the
 Dictyostelium myosin heavy chain.";
 RL FEBS Lett. 227:71-75(1988).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-762.
 RX MEDLINE=95345066; PubMed=7619795;
 RA Fisher A.J., Smith C.A., Thoden J.B., Smith R., Sutoh K., Holden H.M.,
 RA Rayment I.;
 RT "X-ray structures of the myosin motor domain of Dictyostelium
 discoideum complexed with MgADP.Befx and MgADP.ALf4-.";
 RL Biochemistry 34:8960-8972(1995).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 1-762.
 RX MEDLINE=95345067; PubMed=7619796;
 RA Smith C.A., Rayment I.;
 RT "X-ray structure of the magnesium(II)-pyrophosphate complex of the
 truncated head of Dictyostelium discoideum myosin to 2.7-A
 resolution.";
 RL Biochemistry 34:8973-8981(1995).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
 RX MEDLINE=96206189; PubMed=8611530;
 RA Smith C.A., Rayment I.;
 RT "X-ray structure of the magnesium(II).ADP.vanadate complex of the
 Dictyostelium discoideum myosin motor domain to 1.9-A resolution.";
 RL Biochemistry 35:5404-5417(1996).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-762.
 RX MEDLINE=97452580; PubMed=9305951;
 RA Gulick A.M., Bauer C.B., Thoden J.B., Rayment I.;

RT "X-ray structures of the MgADP, MgATPgammAs, and MgAMPPNP complexes
 of the Dictyostelium discoideum myosin motor domain.";
 RL Biochemistry 36:11619-11628(1997).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
 RX MEDLINE=98070605; PubMed=9405148;
 RA Bauer C.B., Kuhlman P.A., Bagshaw C.R., Rayment I.;
 RT "X-ray crystal structure and solution fluorescence characterization
 of Mg.2'(3')-O-(N-methylanthraniloyl) nucleotides bound to the
 Dictyostelium discoideum myosin motor domain.";
 RL J. Mol. Biol. 274:394-407(1997).
 CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE
 ACTIVITY THAT IS ACTIVATED BY ACTIN.
 CC -1- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
 INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI
 LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
 (MLC-2).
 CC -1- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL
 CORTEX.
 CC -1- DOMAIN: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN BE FURTHER
 SPLIT INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 SUBFRAGMENT (S2).
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -1- PTM: PHOSPHORYLATION INHIBITS THICK FILAMENT FORMATION AND REDUCES
 THE ACTIN-ACTIVATED ATPASE ACTIVITY.
 CC -1- MISCELLANEOUS: DICTYOSTELIUM MYOSIN II HAS NO K(2)EDTA ATPASE
 ACTIVITY, PERHAPS CORRELATED WITH THE ABSENCE OF A CYS AT THE SH-1
 POSITION (688).
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 CC EMBL; M14628; AAA33227.1; -;
 DR PIR; A26655; A26655.
 DR PIR; S00250; S00250.
 DR PDB; 1MMA; 03-DEC-97.
 DR PDB; 1MMD; 17-AUG-96.
 DR PDB; 1MMG; 03-DEC-97.
 DR PDB; 1MMN; 03-DEC-97.
 DR PDB; 1MND; 17-AUG-96.
 DR PDB; 1MNE; 17-AUG-96.
 DR PDB; 1VOM; 23-DEC-96.
 DR PDB; 1LVK; 28-JAN-98.
 DR DictyDb; DD01008; mhca.
 DR InterPro; IPR000048; IQ region.
 DR InterPro; IPR004009; Myosin_N.
 DR InterPro; IPR001609; myosin_head.
 DR Pfam; PF00063; myosin_head; 1.
 DR Pfam; PF00612; IQ; 2.
 DR Pfam; PF02736; Myosin_N; 1.
 DR PRINTS; PR00193; MYOSINHEAVY.
 DR PRODom; PD000355; myosin_head; 1.
 DR SMART; SM00015; IQ; 1.
 DR SMART; SM00242; MYSC; 1.
 DR PROSITE; PS50096; IQ; 1.
 DR Myosin; Coiled coil; Actin-binding; ATP-binding; 3D-structure;
 KW Calmodulin-binding; Methylation; Alkylation; Phosphorylation.
 FT DOMAIN 1 761 MYOSIN HEAD-LIKE.
 FT DOMAIN 762 791 IQ.
 FT DOMAIN 817 2116 COILED COIL (POTENTIAL).
 FT NP_BIND 179 186 ATP.
 FT DOMAIN 638 660 ACTIN-BINDING.
 FT DOMAIN 738 752 ACTIN-BINDING.

```
FT MOD_RES 130 130 METHYLATION (DI-) (POTENTIAL).
FT MOD_RES 678 678 ALKYLATION (SH-1).
FT MOD_RES 1823 1823 PHOSPHORYLATION (BY MHCK).
FT MOD_RES 1833 1833 PHOSPHORYLATION (BY MHCK).
FT MOD_RES 2029 2029 PHOSPHORYLATION (BY MHCK).
SQ SEQUENCE 2116 AA; 243871 MW; 2FC3770BBIEE56AI CRC64;

Query Match 11.5%; Score 414; DB 1; Length 2116;
Best Local Similarity 22.3%; Pred. No. 5, 4e-07;
Matches 191; Conservative 178; Mismatches 275; Indels 214; Gaps 36;

QY 8 LKRFNDPSCASPAYD-VKILEVLKGPVSPQK-----SQRFKQKQSKQNLNVDK--- 58
Db 959 MKRVND-----GQSDTRISLEIKD--ELQKEVEELTESFSESKDKGVLEKTRVRL 1008
QY 59 -----DTTLPASARKVKSESKQNDKLDKILEKEIRVLLQBERGAQDRIOQLETEL 113
Db 1009 QSELDLTVRLDSEYKDKSELRLQKKLEELKQVQOEALAAATAKLAQEAANKLQGEY 1068
QY 114 EKMEARLNALREKTSLSANNATLEKQLIELTRTNELLSKSFSENGNOKNRLRILSLELMK 173
Db 1069 TELNEKFNSEVTARSNVESKKTLESQLVAV--NNELDEEKNRDALEKKKKALDAMLEE 1126
QY 174 LNKRETKRMGMKQEGEMK-----LQVTRS-LEESQGKTAQLEGKLVSTEEKIDEK 228
Db 1127 MKDQLES-----TGGEKKSLYDLKVQESDMEALRNQISELQSTIAKLEIKKSTLE 1177
QY 229 SETEKLELYIEISCASDOVERKYK---LDI---AQLPEENLEKENDETLSLKQSLLENI 280
Db 1178 GEVARLOGLEAEQAOLAKSNVEKQKKVELDLEDKSAQAEATAK-QALDKLKKLEQEL 1236
QY 281 VILSKQVEDLVKCOLLEKEKEHDVNRNREHNENLNAEQNLKQFLEQOQHEKLOOKE 340
Db 1237 SEVQOTQLSAANK-----NVNSD-STNKHLETSPNNLKLEAEQKAKQALEKKR 1285
QY 341 LAIDS-----LLOQEKLSLHQKCSFOEMVKENLFEELKQTLDELDKLOOKE 393
Db 1286 LGLESELKHVNLQEEFKQKSEKRVKVDLEKVEYSELKDQIEEVASKAVTEAKNKE 1345
QY 394 EQAERLVKOLEEAKSR---AEELKLEK-----LKGKAELEKSSAAHTQATLLQ 443
Db 1346 SELDEIKROYADVSRDKSVQLTKLQAKNEELRNTEAEABQGLDRAERSKKAEFDLE 1405
QY 444 EKYDSWOSLEDVTA-QFESYKAL-----TASEIEDLK----- 475
Db 1406 E----AVKNLEBETAKVKAEMKKAETDYRSTRKSELDDAKMVSSEYVQIKRNEELS 1461
QY 476 ----LE-----NSLQEK-----AAKAGKNAEDVOHQILATES 504
Db 1462 ELRSVLEEADERCNSAIKAKKTAESALESKDEIDAANNAKAKAPKSKSELEVRVAELE 1521
QY 505 S-----NQEV-----RMLLDLPQKSALKETEIKE----- 529
Db 1522 SLEDSKSGVNVFEIRKDAEIDDLRLDRETESIKSDEDEKKNTRKQFADLEAKVERAQ 1581
QY 530 ----ITVSFLQK-----ITDLQNLQKQE-----EDFRKQLEDE-EGRKAENKTAEIT 574
Db 1582 REWVTIDRLKKLESIDIILSTQLOTETKSRKIEKSKKLEQLTAEARRAEEGSSKAAD 1641
QY 575 EINKRWRLLYELYNKTPFQLODAFEVEKQALLNEHGAQEQNLKTRDSYAKLLGHON 634
Db 1642 EIRK-----QVWQEVDELRAQLDS-----ERAALN---ASEK-KIKSLVAEY---DE 1682
QY 635 LKQKTKHVVKLDE---NSQLKSVSKLRCOLAKKKOSETKLOEELNKVLGIKHFDPK 690
Db 1683 VREQLEDEILAKDLKLVAKRALEVELEEVROQLEEEEDSRSEL-EDSKRRLLTTEVEDIKK 1741
QY 691 APFHESKENFALKTPLE 708
Db 1742 KYDAEVEQNTKLEAKK 1759
```

RESULT 10

Query Match

11.4%; Score 409.5; DB 1; Length 1939;

```
MYH4_HUMAN
ID QY623; STANDARD; PRT; 1939 AA.
AC QY623;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, skeletal muscle, fetal (Myosin heavy chain IIB)
DE (MYHC-IIB).
GN MYH4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=93318869; PubMed=10388558;
RA Weiss A., Schiaffino S., Leinwand L.A.;
RT "Comparative sequence analysis of the complete human sarcomeric myosin
heavy chain family: implications for functional diversity.";
RL J. Mol. Biol. 290:61-75(1999).
CC -!- FUNCTION: MUSCLE CONTRACTION.
CC -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CHARACTER OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
SUBFRAGMENT (S2).
CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
EMBL: AF111783; AAD29949.1; -.
HSP: P13538; 2MYS.
Genew; HGNC:7574; MYH4.
MIN; 160742; -.
InterPro; IPR000048; IQ_region.
InterPro; IPR004009; Myosin_N.
InterPro; IPR002928; Myosin_tail.
InterPro; IPR001609; myosin_head.
Pfam; PF00063; myosin_head; 1.
Pfam; PF00612; IQ; 2.
Pfam; PF01576; Myosin_tail; 1.
Pfam; PF02736; Myosin_N; 1.
PRINTS; PR00193; MYOSINHEAVY.
ProDom; PD000355; myosin_head; 1.
SMART; SM00015; IQ; 1.
SMART; SM00242; MYSC; 1.
PROSITE; PS50096; IQ; 1.
Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
Calmodulin-binding; ATP-binding; Methylation; Alkylation;
Multigene family.
DOMAIN 1 784 MYOSIN HEAD-LIKE.
DOMAIN 785 814 IQ.
DOMAIN 843 1939 COILED COIL (POTENTIAL).
FT NP_BIND 179 186 ATP (POTENTIAL).
FT SEQUENCE 1939 AA; 223012 MW; 40BLAD1D777A47DE CRC64;
SQ
```

Best Local Similarity	22.7%	Pred. No. 6.9e-07;	Matches 186;	Conservativity 131;	Mismatches 233;	Indels 271;	Gaps 30;
QY	136	TLEKQIE----	LTRTNELIKSFSENGKNRIILSLMLKLRNKRET-----	180			
Db	776	TLEMRDEKLAQLTRTQAICRG-----	FLMRVERKMMERRESIFCIQYNIRA	824			
QY	181	-----	KMGMMAKQE-----	GMEKMLQVTOBSLEESOCKIAQLSGKLVS	220		
Db	825	FMNVKHPWPKLYFKIPKLLKSAETEKEMANKEFEKTBELAKTEAKRKELEBEKVMVL	884				
QY	221	EKKIDKSETEKLELVEIEBISCASDOVEKYKL-----	DIAOLEENLKEKNDEILSLK	273			
Db	885	MOEKNDLOLOVQADALAAAEERCDOLIKTKIOLEAKIKEVTEBRAEDEEINAEELTAKK	944				
QY	274	QSLSEENTVIILSKOVEDLNVCQKLEKEDEHVRNREHNENLNEMQNL-----	KOK	325			
Db	945	RKLEDECSSELKKDIDLELTILAKVEKEK--	HATENKV--KNLTEMAGLDETIAKLPIKEK	1000			
QY	326	FILB---QOBEHEKQKQELQIDSL-----	QOBEKELSSSLHO--KLCSQOE-----	366			
Db	1001	KALQEAHQOTLDDIQEEDKVTNTTKATKLEQQVDDLEGSLEQEKKLCMDLERAKRKLE	1060				
QY	367	---EMVKEKNLFEELKOTLDELDKLOOKE-----	EQA-----	396			
Db	1061	GDUKLAQESTMTDENDKOOLNE--	KLKKKEFEMSNLOCKIEDEQALAMQLQKKIKELQAR	1118			
QY	397	ERLVKOLEEEAEKRAE-----	ELKLEELK-----	427			
Db	1119	IEELEIEIEARSAKAEKORSDLSELEIEISRLPEAGGATSAQIELNKKRAEFQKM	1178				
QY	428	---LEKSSAHTQATLLQEKY-DSM-----	VOSLEDVTAQFESYKALTASEIEDLKLE	477			
Db	1179	RRDLEESTLQHEATAAALRKKKHADSVAEGLQKQIDSLQKVQKLEKSEKSELKWEINDLASN	1238				
QY	478	NSSLQEKAAKAGNAEDVQHOILATESNQBYVRMLDLQTKSALKETEIKEIT-----	531				
Db	1239	METVSKAKANFEKRCRTLEQDLSIKTKEEQORQLINELSAQKARLHTESEGEFSRQDLEK	1298				
QY	532	---VSFLOKITDLONLQOKEED-----	FRQLDEDEGRK	563			
Db	1299	DAMVSQLSRGKQATQOIEELKQLEETKAKSTLAHALQSAHRDCLLREQYEEEQAK	1358				
QY	564	AENKTAAELTEENKVRLLY-----	EELYNKTRPFQLOLD-----	601			
Db	1359	AELQGRGSKANSEVAQNRKTYETDAIQITTELEAKKLAQLODAEHEVAVNSKCSAL	1418				
QY	602	EVEQKALLNE-----	HGAQEOINKTROSAYAKLLGHONLK-----	636			
Db	1419	EKTQRLQNEVEDLMIDVERNSNAACIALDKQRNFQKVLAEWKQYETQAELEASOKES	1478				
QY	637	-----	QKIKHVVLKDNSOLKSVEVSKLRCOLAKKQSEKTLQELKNVL	681			
Db	1479	RSLSLELTKVKNAYEESLDHLETLKRNKNLQOESDITQIAE-----	GGKHIHELEKV--	1533			
QY	682	GIRKFDPSKAFHESKNFALKTPLKRGENTNCRYAPMECQE	722				
Db	1534	-----KKQLDHEKSE--	LOTSLEA-----	EASLEHEE	1559		
RESULT 11							
MYH9_CHICK							
ID	MYH9_CHICK	STANDARD;	PRT;	1959	AA.		
AC	P14105;						
DT	01-JAN-1990 (Rel. 13, Created)						
DT	01-JAN-1990 (Rel. 13, Last sequence update)						
DT	16-OCT-2001 (Rel. 40, Last annotation update)						
DE	Myosin heavy chain, nonmuscle (Cellular myosin heavy chain) (NMHC).						
GN	MYH9.						
OS	Gallus gallus (Chicken).						
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianinae;						
OC	Gallus.						

Qy	202	SLE	-----ESQKIA-----	QLEGGKLVSTEKE-----	KIDE-KS	229
Db	1052	KLEGGSDLDHQIAELQAOIAELKIQLSKKEEELQAALRAVEEAAQNALMKAKIRELS	1111	1111	1111	1111
Qy	230	ETEKLLEYIEISCASDQVEKYKLDIAOLEENLKEKNDIL-----	SLKOSLEENIVI	282		
Db	1112	QITELQEDLESERASNKAQKQKRLGEELEALKTELEDITLDSTAAQOELSKREQEVTV	1171	1171	1171	1171
Qy	283	LSKQVED-----	LNVRQCQLLEKEKEDHVN	306		
Db	1172	LKKTLEDEAKTHEAQIQEMRQKHSQAIBELAQESQTKRVRKANLEKAKQALESERAELSN	1231	1231	1231	1231
Qy	307	R-----NRH-NENINAEQMLKOKFI-----	LEQOEHEKIQOELQIDS-----	LLQO	349	
Db	1232	EVKVLQGGKGDAAHEKRRKVDQAOLQELQVKKFTTEGERVKTAEARVKNLQVLQEDNVTGLNQ	1291	1291	1291	1291
Qy	350	EKELS-----SSLHKQLCSQEEEMVAKNL-----	FEELKQTLDELDKLOOEQEARL	399		
Db	1292	SDSKSTKLAKDFSALBSQIQDTQELLQBETRLKLSFSFKLKQTEDEKNALKEQLEEEEA	1351	1351	1351	1351
Qy	400	VKQLEE-----EAKSRAEE-----	LKLEELKELKGAELAELEKSSAAHTQ-----	ATLLL	442	
Db	1352	KRNLEKQISVLQOAAVEARKWDDGGLCLETAAEAKKQLDKLDESILQRTVEEKTAAVDKL	1411	1411	1411	1411
Qy	443	QEKYDSNMVOSLEDVTAQFESYKALTASETE-----	DLKLENSLSQEKAA-----	RAGKN	491	
Db	1412	EKTTRKLQELDDIAVDLD-HQRTQVSNLEKKKQKFDQLLAAEKNISAKYAEERDRAEE	1470	1470	1470	1470
Qy	492	AEDVQHQILATESNQEYVRMLLDIQTSAKETEIKEITVS-----	FLQKITDQLQNL	545		
Db	1471	AREKETKALSARALEAEATEQAELERVNKQFRTEMEDLMSSKDDVGSVHELEKAKRAL	1530	1530	1530	1530
Qy	546	KQOEBDFRQLEDGEGRKAEKENTTAELTEETENKWR-----	LLYEELYNKTKPFOL-	596		
Db	1531	EQOEVEMKQTLQEELEDELQATED-AKULEVQNQAAMKAFQDRDLGRDQONEEKKRQLI	1588	1588	1588	1588
Qy	597	QIDAFEFVEKQALLNHHGAABQLAKIRDSYAKLLGH-----	QNLKQIKHVKVLKLDENS	650		
Db	1589	RQVREMEVELEDERKQSTAVAAARKKLELDLKLQLESHDITANKNRDEAIKHVRKQLQAOMK	1648	1648	1648	1648
Qy	651	QKASEVSKLRQ-----LAKKQOSETK-----	LOEEL	677		
Db	1649	DYMRELEDRTSGRETLAQAKENKKKLSMEAMIQLOEEL	1689	1689	1689	1689

RESULT 12

MYSN_DROME

AC MYSN_DROME STANDARD; PRT; 2017 AA.

AD Q99323;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Myosin heavy chain, non-muscle (zipper protein) (Myosin II).

GN ZIP.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;

OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_Taxid=722;

RP [1]

RN SEQUENCE FROM N.A.

RC TISSUE=Embryo;

RC MEDLINE=90349606; PubMed=2117279;

RA Ketchum A.S., Stewart C.T., Stewart M., Kiehart D.P.;

RT "Complete sequence of the Drosophila nonmuscle myosin heavy-chain transcript: conserved sequences in the myosin tail and differential splicing in the 5' untranslated sequence.";

RL Proc. Natl. Acad. Sci. U.S.A. 87:6316-6320(1990).

CC -1- FUNCTION: NONMUSCLE MYOSIN APPEARS TO BE RESPONSIBLE FOR CELLULARIZATION. REQUIRED FOR MORPHOGENESIS AND CYTOKINESIS.

CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.

CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.

CC	-!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL: M35012; AAA28713.1; -
DR	PIR: A36014; A36014.
DR	PIR: B36014; B36014.
DR	HSSP: P10587; IBR2.
DR	FLYBase: FBgn005634; zip.
DR	InterPro: IPR000048; IQ_region.
DR	InterPro: IPR004009; Myosin_N.
DR	InterPro: IPR002928; Myosin_tail.
DR	InterPro: IPR001609; myosin_head.
DR	Pfam: PF00063; myosin_head; 1.
DR	Pfam: PF00612; IQ; 1.
DR	Pfam: PF01576; Myosin_tail; 1.
DR	Pfam: PF02736; Myosin_N; 1.
DR	PRINTS: PR00193; MYOSINHEAVY.
DR	ProDom: PD000355; myosin_head; 1.
DR	SMART: SM00015; IQ; 1.
DR	SMART: SM00242; MYSC; 1.
DR	PROSITE: PS50096; IQ; 1.
KW	Myosin; Alternative splicing; Coiled coil; Actin-binding;
KW	ATP-binding; Calmodulin-binding.
FT	DOMAIN 1 829 MYOSIN HEAD-LIKE.
FT	DOMAIN 830 859 IQ.
FT	DOMAIN 886 2017 COILED COIL (POTENTIAL).
FT	NP_BIND 225 232 ATP.
FT	DOMAIN 250 260 25 KDA/50 KDA JUNCTION.
FT	DOMAIN 682 694 50 KDA/20 KDA JUNCTION.
FT	DOMAIN 705 727 ACTIN-BINDING.
FT	DOMAIN 742 758 REACTIVE SULFHYDRYL/ACTIN-BINDING.
FT	DOMAIN 1303 2017 LIGHT MEROMYOSIN (LMM).
FT	DOMAIN 1303 1970 ALPHA-HELICAL TAILPIECE (LMM).
FT	DOMAIN 1971 2017 GLOBULAR TAILPIECE.
FT	VARSPLITC 1 45 MISSING (IN SHORT ISOFORM).
SQ	SEQUENCE 2017 AA; 232016 MW; 73E3CH02BA8F2528 CRC64;
	Query Match 11.3%; Score 408; DB 1; Length 2017;
	Best Local Similarity 24.7%; Pred. No. 8e-07;
	Matches 209; Conservative 134; Mismatches 286; Indels 218; Gaps 35;
Qy	29 LEVLKGPVSF-QKSQRFKOOKESQNLN-----VDKDT-----TLPA 64
Db	889 LEVTKQEKLVQDELDKQVREKLDLAKNTQEVRYQQALVEKTTLAELQAEIELCA 948
Qy	65 SARVKVS-----SESKESQNKDLKTLKEIRVLL--QERGAQDRRIQDLETELEKMEA 118
Db	949 EAEESRRLMARKQEDMDMQELETREEEERYVALLGEGKKLELNIQDLEEQLEEEEA 1008
Qy	119 RLNAALREKTSLSA-----NNATI-EKQLIELTRTNELLAKSKFSENGNQK 162
Db	1009 AROKLQLEKVLQDAIKKYEDIALTDQKLLKEKKLLLE-ERANDLSQTLAEEBEAK 1067
Qy	163 NLRILSLEMLKRNKRE--TKMRGMMAKGEMKMLQVTRSLSE-----ESQG 208
Db	1068 H-----LAKLRKHEATITEELERLHKDQQQESDRSKRKTIETEVADLKQQLNERY 1120
Qy	209 KIAQLEKGLVSIKE-----KIDKSETE-----KLEFYIEISCASQD 247
Db	1121 QVDEMQAQLAKREBELTQTLLRIDEESATKATAQKAQRELSQLAEIQDLEAKAARAK 1180
Qy	248 VEYKDKDIALEENLK-----EKNDIELSKOSLEENIV-----281
Db	1181 AEKVRRLDSEELBAKKNELDLSLQTTAAQQQLSKRQELATLUKKSLEETVNHGVLAD 1240
Qy	282 -----ILSKQVEDLVKQQLLEKEKED-----HVNRRNREHNELN-- 316

```

Query Match      11.3%; Score 408; DB 1; Length 2017;
Best Local Similarity 24.7%; Pred. No. 8e-07;
Matches 209; Conservative 134; Mismatches 286; Indels 218; Gaps 35;

Qy 29 LEVLKGPVSF-QKSQRFKQOKESKQNUN-----VDKDT-----TLPA 64
    |||  ||  |||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 889 LEVTKQEKLVQKDELVKQVREKLDLTAKNTQEVERKYQQALVEKTTLAEQLQAEIELCA 948

Qy 65 SARKVKS-----SESKKSQNNDLKLLEKIRVLL--QERGQADRRIQDLETELEKMEA 118
    :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 949 EAEESRRLMARKQEDMMQOELETRIEEEERVLALGGEKKKLELNLQDLEQLEBEEA 1008

Qy 119 RLNALREKTSLSA-----NNATL-EKQILERTNELLKSKSENGCOK 162
    ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 1009 ARQKLQLEKVLQDAKIKKYEEEDLALTDQNKLLKEKKLLE-ERANDLSOTLAEEBEAK 1067

Qy 163 NLRILSLEMLKLNKRE--TKRGMMAKQEGMEMKLOVTTQRSLE-----ESQG 208
    :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 1068 H-----LAKLAKHAEATTILEERLHKDQOQQRQESDRSKRKRIETEVADLKEQLNERRY 1120

Qy 209 KIAQLEGLKVSIEKE-----KIDKSETE-----KLLEYIEBISCASDQ 247
    :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 1121 QVDEMQAQLAKRBEELTQTLRLRIDEESATKATQAQAQRELESQALAEQBLEAKAKARAK 1180

Qy 248 VERKYDLIAOLEENLK-----EKNDLEILSKQSLEENIV-----281
    ||  :  :  :  :  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 1181 AEKVRDRLSEELFALKNELLDLSLDTTAAQOELSKRQEQLATLUKSLSEETVNHGVGLAD 1240

Qy 282 -----ILSKQVEDLVNKKQLEKEKED-----HVNRRNRHENENLN-- 316

```


QY 430 KSSAHTQA-----TLLQKQYDSWQSLDVT-----QFESYKALTAETE 472
Db 821 KNLAVNQKDSLEKELQKKEFTSAVDAENARQAMQETINKLNKEEQFALMSSELE 880
QY 473 DLKLENSLOEKAARAKAGNAED-----VOHQILATESSNOEYVRLMLLDLOTSAKLE 524
Db 881 QLK-SNLTYMETKLEREREQOLTEAKYKLENDIAETIMKSSGSSAQLMKMNDLRLKE 939
QY 525 TEIKETVSFL--OKITDLQNLQKQ-----QEE-----DPRKOL 556
Db 940 RLEQIQLELTANKEAVOLQKNVQTAQKABQSOQETIKTHQEEELKMKQDQLOTKMKOM 999
QY 557 EDEEGR---KAEKENTTAELTEEN-----KWRLLYEELYNKT---KPFQQLQDAFEV 603
Db 1000 ETSNQYKDLQAKYKETSEMITKHADIKGFKQNLDAEALKAQKKNDELETAQAEEL 1059
QY 604 EKQALLNEHGAQEQOLKIRDSYAKLLGHQNL--KQKTHVVKLKDENSQKSEVSKURC 661
Db 1060 KQQA---EQAKADKRAEEVLOTMERVTRKDAIHOEKIETTLASL--ENSROTN--KLQN 1112
QY 662 OLAKKKQSETKLOEELNVLGKTHDPSPKAFHESKENF-ALK 703
Db 1113 ELDMLKQNNLKNKEBELTKSKELLNENKV--BELKKEFEALK 1153

RESULT 14
MYH2_HUMAN
ID MYH2_HUMAN STANDARD; PRT; 1941 AA.
AC Q9UHX2: Q16229; Q14322;
DT 16-OCT-2001 (rel. 40, Created)
DT 16-OCT-2001 (rel. 40, Last sequence update)
DT 16-OCT-2001 (rel. 40, Last annotation update)
DE Myosin heavy chain, skeletal muscle, adult 2 (Myosin heavy chain Iia)
DE (MyHC-Iia).
GN MYH2 OR MYHSA2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=99318869; PubMed=10388558;
RA Weiss A., Schiaffino S., Leinwand L.A.;
RT "Comparative sequence analysis of the complete human sarcomeric myosin heavy chain family: implications for functional diversity";
RL J. Mol. Biol. 290:61-75(1999).
RN [2]
RP SEQUENCE OF 1711-1941 FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=95109625; PubMed=7545970;
RA Smerdu V., Karsch-Mizrachi I., Campione M., Leinwand L.,
RA Schiaffino S.;
RT "Type Iix myosin heavy chain transcripts are expressed in type Iib fibers of human skeletal muscle";
RL Am. J. Physiol. 267:C1723-C1728(1994).
RN [3]
RP SEQUENCE OF 1823-1941 FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=95270723; PubMed=7751403;
RA Enlioni S., Sant'ana Pereira J., Sargeant T., Young A., Goldspink G.;
RT "Characterization of human skeletal muscle fibres according to the myosin heavy chains they express";
RL J. Muscle Res. Cell Motil. 16:35-43(1995).
CC -1- FUNCTION: MUSCLE CONTRACTION.
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES, CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY

CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED SUBFRAGMENT (S2).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC EMBL; AF111784; AAC29950.1; -.
CC EMBL; S73840; AAC13916.1; -.
CC EMBL; P32858; CAAB3687.1; -.
CC HSSP; P13538; 2MVS.
CC Genew; HGNC:7572; MYH2.
CC MIM; 160740; -.
CC InterPro; IPR000048; IQ_region.
CC InterPro; IPR004009; Myosin_N.
CC InterPro; IPR002928; Myosin_tail.
CC InterPro; IPR001609; myosin_head.
CC Pfam; PF00063; myosin_head; 1.
CC Pfam; PF00612; IQ; 2.
CC Pfam; PF01576; Myosin_tail; 1.
CC Pfam; PF02736; Myosin_N; 1.
CC PRINTS; PR00193; MYOSINHEAVY.
CC ProDom; PD000355; myosin_head; 1.
CC SMART; SM00015; IQ; 1.
CC SMART; SM00242; MYSC; 1.
CC PROSITE; PS50096; IQ; 1.
CC Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding; Calmodulin-binding; ATP-binding; Methylation; Alkylation;
KW Multigene family.
KW DOMAIN 1 786 MYOSIN HEAD-LIKE.
FT DOMAIN 787 816 IQ.
FT DOMAIN 845 1941 COILED COIL (POTENTIAL).
FT NP_BIND 179 186 ATP (POTENTIAL).
FT CONFLICT 1844 1844 K -> R (IN REF. 3).
SQ SEQUENCE 1941 AA; 223043 MW; 681E866F83AE83F CRC64;
Query Match 11.38; Score 407; DB 1; Length 1941;
Best Local Similarity 24.08; Pred. No. 8.3e-07;
Matches 211; Conservative 136; Mismatches 253; Indels 278; Gaps 39;
Qy 51 KQNLNVKDTTLPASARKVSKESKESQKNDKDLKILEKIRV-----LLQ 97
Db 725 KQRYVLNNSAIPG---QFIDSKKASEKLLASIDIDHTQYKFGHTKVFVKAGLLGLE 780
Qy 98 ERGAQDRRIQDLETELEKMEARLNALREKTSLSANNATLEKQIETRTNELLKSKFSE 157
Db 781 E--MRDDKLAQLIT---RTQARCRGFLAR-----VEYQRMVERREAFICI 820
Qy 158 NGNKN-LRILSLEMLKLNKRETKMGMMAKQGMK--LQVTSRLESQKIAOLE 214
Db 821 QYNTRFMNVKHPWPKLFFKIPKLLKSAETKEKEMATMKKEEFQIKDELAKSEAKRKELE 880
Qy 215 GLKYSIEKEIDKSETEKLELEYIEEISCASDQVEKYK-----DIAQLEENLKEKND 267
Db 881 EKMYTLLKEKNDLQVQAEAGLADAEERCDDLIKTIQLEAKIKVEYTERADEEEINA 940
Qy 268 EILSKOSLENIIVLSKQVEDLNKCOLLEKEKEDHVNREHNENLNAMQN-----322
Db 941 ELTAKKRLEDECSLAKKDIDDLTLLAKVEKEK--HATENKV--KNITEMAGLDEITIA 996
Qy 323 ---KQKFTLE-----QOEHEKIQ-----QKQLQIDSL---LQOEKELSSSLHQ 359
Db 997 KLTRKALQEAHQOITLDDLAQEDKVNLTAKIKLEQQVDDLEGSLEQEKKLMDLER 1056

QY 360 -----KLCFQEMV---KEKNLFEELKQTLDELKLOOK---EEQA----- 396
Db 1057 AKRLEGLKLA--QESIMDIENKQOLDEKLKKEFEISNLQSKIEDQALGIQLOKKI 1114
QY 397 -----ERLVKOLEPEAKSRAE-----ELKLLLEKL-----KGRE 425
Db 1115 KELQARIELEEIEAEASRAKAEKORSDSLSELEIEISERLEBEAGGATSAQIEMNKKRE 1174
QY 426 AE-----LEKSSAHTQATLLQEKY--DSM-----VQSLBDVTAQFESYKALTASEI 471
Db 1175 AEPQKMRDLLEATLQHEATAATLRKKHADSVAEELGEQIDNLRQVKQLEKEKSEMKMEI 1234
QY 472 EDLKLNESSLOEKAAGKAGNAE----DVQHILATESSNQEYVRMLLD-----LQTKS 520
Db 1235 DDL-----ASNVTVSRAKGNLEKMCRTLEDQLSLKSEKEEQORLINDLTAQRGRLOTES 1290
QY 521 A--LKETEKEITVS-----FLQKITDLOQLKQO-----EED---FR 553
Db 1291 GEFSQLDEKEALVSQSRGKQAFQTOOIEELKQLEEEIKAKNALAHALQSSRHDCDLLR 1350
QY 554 KQLEDEEGRAKAKENTTAELTEIEINKWRLLY-----EELYNKTKPFQOLQDA----- 600
Db 1351 BOYEEESKAEALQALRSKANTEVAQWRKYETDARTQTEELEBAKKLAORLOAAEEHV 1410
QY 601 -----FEVEKQALLNE-----HGAAQPOLNKRIRDSYAKLLG-----H 632
Db 1411 EAVNAKASLEKTKQRLQNEVEDLMDVERTNAACAALDKKORNFCKILAENKQKCEETH 1470
QY 633 QNLKQ-----KIKHVVK-----LKENSOLKSEVSKRLCOLAK----- 665
Db 1471 AELEASOKEARSGLTELFKIKNAYESLDQLETLRKENKNLQOEISDLTEQIAEGGKRIH 1530
QY 666 ----KKQ-----SETKLOEELNKVLGK 684
Db 1531 ELEKIKQVEQKCELOQAALAEAEASLEHEEGKILRIQ 1568

RESULT 15
Y109 YEAST
ID Y109 YEAST STANDARD; PRT; 1679 AA.
AC P40457;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Hypothetical 195.1 kDa protein in DNA43-UBI1 intergenic region.
GN Y1149C
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RA Barrell B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,
RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
RA Walsh S.V., Whitehead S.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z38059; CAA86129.1; --
DR F1R; S48385; S48385.
DR SGD; S0001411; MLP2.
KW Hypothetical protein.

SQ SEQUENCE 1679 AA; 195141 MW; 298950CC52202D8F CRC64;
Query Match 11.3%; Score 405; DB 1: Length 1679;
Best Local Similarity 21.9%; Pred. No. 8.5e-07;
Matches 200; Conservative 167; Mismatches 321; Indels 224; Gaps 32;
QY 25 DVKTLEVLKGPVFSQRFKQOKESQNLNVDKDTTLPASARKVSKSSKESQKNKD 84
Db 471 NTSAIQETASPLSQDELISLRKILESSNIVN-ENDSQAIITERLVFESVNLQEKV 529
QY 85 LKILEKEIRVLQOE---RGAQDRRTODLET-----ELE-----KWEARLNAALR 125
Db 530 LNC-----IRLADKLENYEGKDKTLQKVENQTIKEAKDAIIELEINAKMETRIN 585
QY 126 EK-----TSLSANNATLEKOLIELTRTNELLKSKSPSENGKNLRLSL 169
Db 586 ERDSYKLLASTEENKANTNSVTWEAAAREKKIRELEA--ELSSTKVENSAILQNLRELL 643
QY 170 ELMKLRNKRETKM-----RGMMAKQEGM-----EMKLQVTQR----- 201
Db 644 IYKKSQCKKTKTTLEDENFKGLAKERMLLEEAIDHLKAELEKQKSWPSYIVHVEKERAS 703
QY 202 -SLEESQGTQAQLEGLVSTIEKEKID-----EKSTELKLEYIEIESC 243
Db 704 TELSOSRIKIKSLEYEIEIKLKKETASFIPTKESLTRDFEQCCKEKKELQMLRKE-SEISH 762
QY 244 ASDQVE-----KYKLDIAOLEN-----LKEKNDEILSKOSLEENIVILSKOVEDL 290
Db 763 NENKMPFSKEGQTKAKIKLENNLERLSDLQSKIOEISIRSCDSQKWAQNTIDDT 822
QY 291 NVRCOLLEKRP---KEDHVNRRNHNENLNAEMQNLK--QKFI-----LEQOEHEKLQ 337
Db 823 EMKMSKLTTELSNKEITIEKLSSEIENLDKELRKTQYKFLDQNSDASTLEPTLRKELE 882
QY 338 QKELQI---DSLLOQEKELSSLSLHKLCSFQDEMVMKKNLP-----EELK 380
Db 883 QIQVOLKDANSQIQAYEEIISSENNALIELKNELAKTKENYDAKIELEKKEKWAREEDLS 942
QY 381 QTLDEL-----DKLQKEEQEARLVKOLE----- 404
Db 943 RLKELGETRALQPKLKEGALHFVQOSEKURNEVERIQRMIEKIEKMTIVQLCKKEMS 1002
QY 405 -----EAKSRAEELKLEELKLEKBAELEKSSAAHTQATLLQEKYDSMVQSLQEDVTA 458
Db 1003 QYQSTMKENKDLSELVIRLEKDAADCAQELTKTKSSLYSA---QDLDRKHERKWMEEKA 1058
QY 459 QFESYKALPTASEITDKLENSLOEAKAAGNAEDVOHOILATESSNQEYVRMLLDIQT 518
Db 1059 DYERELISNTEQTESLRVENSVLIEKVDDTAANNQGRDKHLKLVSLFSLNLRHNSLETKL 1118
QY 519 KSALKETE-IKEITVSLQKITDLOQLKQOEEDFR-KOLEDEEGRAKAKENTTAELTEE 576
Db 1119 TTCRRELAFYKQKNDLSLEKTIINDLQRTQTUSEKEYQCSAVIIDFKDITKEVTVQNLKE 1178
QY 577 IN-----KWRLLYEELYNKTKPF-QQLQDAFEVEKQALLNEHG--AAEQELNK 621
Db 1179 NNAILQSLKANVTKEKNEIYKQNDROEISRLQORDLIQTKQVSIKNSKILYVESEMEQ 1238
QY 622 IRDSYAKLLGHQNLKQKIKHVVKLKDENSOLKSEVS-----KLRQLAKK- 666
Db 1239 CKQRYQDLSQQQDAQK-KDIEKLTNEISDLKGLSSAENANADLENKFNRLKQAHEKL 1297
QY 667 ---KQSETKLOEELNKVLGK-----HFDPSKAFHESK---ENFALTKPLKEGWTNC 713
Db 1298 DASKKQOAAATNELNELKAIKDLQDLHFENAKVILDLDTKLKAHELQSEDSVNDHEKDT 1357
QY 714 YRAPMECOESWK 725
Db 1358 YRTLMEEIESLK 1369

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:09:23 ; Search time 19.7824 Seconds
(without alignments)
3523.197 Million cell updates/sec

Title: US-09-685-010-47

Perfect score: 3598

Sequence: 1 MSFPKAPLKRFPNDPSCGAPS.....LKEGNTNCRAPMECQESWK 725

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR-73.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3598	100.0	725	1 JC5016	hyaluronan recepto
2	1952	54.3	631	2 JC4298	hyaluronan recepto
3	444	12.3	1790	2 S67593	transport protein
4	442	12.3	2139	2 T18296	myosin heavy chain
5	436.5	12.1	1690	2 T13030	microtubule bindin
6	425.5	11.8	1410	1 A57013	early endosome ant
7	417	11.6	1875	2 S38173	myosin-like protei
8	416	11.6	1738	2 T14867	interaptin - slime
9	414	11.5	1776	2 A59252	myosin heavy chain
10	414	11.5	2116	2 A26655	myosin heavy chain
11	409	11.4	1959	1 A33977	myosin heavy chain
12	409	11.4	2007	1 B43402	myosin heavy chain
13	408	11.3	2017	1 A36014	myosin heavy chain
14	408	11.3	2057	2 S61477	myosin II heavy ch
15	405.5	11.3	1964	2 A59282	nonmuscle myosin I
16	405	11.3	1679	2 S48385	hypothetical prote
17	405	11.3	1961	1 A61231	myosin heavy chain
18	403	11.2	1938	2 A59293	skeletal myosin he
19	403	11.2	1940	1 A24922	myosin heavy chain
20	402.5	11.2	1972	1 A41604	myosin heavy chain
21	402	11.2	1979	1 S03166	myosin heavy chain
22	400	11.1	1999	1 S21801	myosin heavy chain
23	399.5	11.1	1538	2 T29095	cardiac muscle fac
24	399.5	11.1	1938	1 JX0178	myosin heavy chain
25	399	11.1	1931	2 A59234	slow myosin heavy
26	398	11.1	1269	2 F84730	probable myosin he
27	398	11.1	1938	2 JC5421	smooth muscle myos
28	398	11.1	1940	1 S04090	myosin heavy chain
29	398	11.1	1972	2 JC5420	smooth muscle myos

ALIGNMENTS

RESULT 1

JC5016

hyaluronan receptor - human

C:Species: Homo sapiens (man)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 29-Sep-1999

C:Accession: JC5016

R:Wang, C.-J., Entwistle, J.; Hou, G.; Li, Q.; Turley, E.A.

(Gene 174, 299-306, 1996)

A:Title: The characterization of a human RHAMM cDNA: Conservation of the hyaluronan-b

A:Reference number: JC5016; MUID:97045829; PMID:8890751

A:Contents: breast

A:Accession: JC5016

A:Molecule type: mRNA

A:Residues: 1-725 <WAN>

A:Cross-references: GB:U29343

A:Note: it is uncertain whether Met-1 or Met-196 is the initiator

C:Comment: This receptor regulates focal adhesion turnover, and regulates human breas

C:Genetics:

A:Gene: GDB:HMWR; RHAMM

A:Cross-references: GDB:683209; OMIM:600936

A:Map position: 5q33.2-5qter

C:Superfamily: hyaluronan receptor

Query Match	100.0%	Score	3598	DB 1;	Length	725;	
Best Local Similarity	100.0%	Pred. No.	3e-113;				
Matches	725;	Conservative	0;	Mismatches	0;	Gaps	0;

Qy	1	MSFPKAPLKRFPNDPSCGAPSPGAYDVKTLEVLKGPVSFQSKQRFKQKESQNLNVDRDT	60
----	---	---	----

Db	1	MSFPKAPLKRFPNDPSCGAPSPGAYDVKTLEVLKGPVSFQSKQRFKQKESQNLNVDRDT	60
----	---	---	----

Qy	61	TLPASARKVKSSEKSKQKNDKDLKILEKEIRVLLQERGAQDRRIQDLETELEKMEARL	120
----	----	--	-----

Db	61	TLPASARKVKSSEKSKQKNDKDLKILEKEIRVLLQERGAQDRRIQDLETELEKMEARL	120
----	----	--	-----

Qy	121	NAALREKTSLSANNATLPEKLIETRTNELLKSKFSENGNKNLRIILSLELMKLNKRKET	180
----	-----	---	-----

Db	121	NAALREKTSLSANNATLPEKLIETRTNELLKSKFSENGNKNLRIILSLELMKLNKRKET	180
----	-----	---	-----

Qy	181	KMRGMAKQEGMEKMLQVTRSLSEESQGIQALEGKLVISIEKEIDSEKSEKLELEYEE	240
----	-----	---	-----

Db	181	KMRGMAKQEGMEKMLQVTRSLSEESQGIQALEGKLVISIEKEIDSEKSEKLELEYEE	240
----	-----	---	-----

Qy	241	ISCSADQVEKYKLDIAQLEENLKEKNDEILSKQSLEENIVILSKQVEDNLVKKCOLLEKE	300
----	-----	--	-----

Db	241	ISCSADQVEKYKLDIAQLEENLKEKNDEILSKQSLEENIVILSKQVEDNLVKKCOLLEKE	300
----	-----	--	-----

Qy	301	KEDHVRNRREHNENLNAEQNLKQKFILEQOEHEKLOQKEQLQIDSLLOEQEELSSSLHQK	360
----	-----	--	-----

Db	301	KEDHVRNRREHNENLNAEQNLKQKFILEQOEHEKLOQKEQLQIDSLLOEQEELSSSLHQK	360
----	-----	--	-----

Qy	361	LCSFQEEVVKENLFEELKQTLDELDKLOQKEQAERLVKQLEBEAKSRAEELKLLEEK	420
----	-----	---	-----

Db	361	LCSFQEEVVKENLFEELKQTLDELDKLOQKEQAERLVKQLEBEAKSRAEELKLLEEK	420
----	-----	---	-----

|||||
Db 361 LCSPFQEMVKEKNLFEELKOTLDLQKEQAEALYKOLEBEAKSABELKLLBEK 420
Qy 421 LAGKAELEKSSAAHTQATLLQEQYDSMVQSLDVTQAFESYKALTASEIEDLKLENS 480
Db 421 LAGKAELEKSSAAHTQATLLQEQYDSMVQSLDVTQAFESYKALTASEIEDLKLENS 480
Qy 481 LOEKAAGKAGNAEDVQHQILATESSNOEYVRMLDLQTKSALKETEIKETIVSFQKITYD 540
Db 481 LOEKAAGKAGNAEDVQHQILATESSNOEYVRMLDLQTKSALKETEIKETIVSFQKITYD 540
Qy 541 LQNLQKQBEDPRKQLEDEBGRKAKKENTTAELTEINKWRLLYBELYNKTKPFQQLDA 600
Db 541 LQNLQKQBEDPRKQLEDEBGRKAKKENTTAELTEINKWRLLYBELYNKTKPFQQLDA 600
Qy 601 FEVEKQALLNEHGAEOELNKRDSYAKLLGHONLKQKIKHVVKLKDENSQKSEVSKLR 660
Db 601 FEVEKQALLNEHGAEOELNKRDSYAKLLGHONLKQKIKHVVKLKDENSQKSEVSKLR 660
Qy 661 COLAKKKQSETKLQBELNKNVLGKHFDPSPKAPHHESKENFALTKPLKEGNTNCRYAPMEC 720
Db 661 COLAKKKQSETKLQBELNKNVLGKHFDPSPKAPHHESKENFALTKPLKEGNTNCRYAPMEC 720
Qy 721 QESWK 725
Db 721 QESWK 725
RESULT 2
JC4298
hyaluronan receptor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 16-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 01-Dec-2000
C:Accession: JC4298; A42925; A41923; S21586
R:Enwestle, J.; Zhang, S.; Yang, B.; Wong, C.; Li, Q.; Hall, C.L.; A, J.; Mowat, M.; G
Gene 163, 233-238, 1995
A>Title: Characterization of the murine gene encoding the hyaluronan receptor RHAMM.
A:Reference number: JC4298; MUID:96011639; PMID:7590272
A:Accession: JC4298
A:Molecule type: mRNA
A:Residues: 1-631 <ENT>
A:CROSS-references: EMBL:X64550
A:Experimental source: 3T3 fibroblast
R:Hardwick, C.
J. Cell Biol. 118, 753, 1992
A:Reference number: A42925; MUID:92348516; PMID:1639856
A:Contents: erratum
A:Accession: A42925
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 155-376,'S',378-504,'E',506-631 <HAR>
A:CROSS-references: GB:X64550
A>Note: authors translated the codon AGT for residue 377 as Thr and CTA for residue 507
R:Hardwick, C.; Hoare, K.; Owens, R.; Hohn, H.P.; Hook, M.; Moore, D.; Cripps, V.; Auste
J. Cell Biol. 117, 1343-1350, 1992
A>Title: Molecular cloning of a novel hyaluronan receptor that mediates tumor cell motil
A:Reference number: A41923; MUID:92299690; PMID:1376732
A:Accession: A41923
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 153-376,'S',378-504,'E',506-507,'L',508-630 <HA2>
A:CROSS-references: GB:X64550
A>Note: this sequence has been corrected in reference A42925
C:Comment: This protein regulates cell motility and transformation, and focal adhesion d
C:Genetics:
A:Gene: rhamm
A:Introns: 20/3; 54/2; 79/2; 137/3; 187/3; 259/2; 382/2; 431/2; 482/2; 515/3; 574/3; 625
C:Superfamily: hyaluronan receptor
C:Keywords: glycoprotein; receptor
F:260-382/Region: 21 residue repeats
F:516-574/Region: hyaluronan binding #status predicted
F:575-625/Region: hyaluronan binding #status predicted
F:116,283,304,325,346,367,398,438,619/Binding site: carbohydrate (Asn) (covalent) #statu

Query Match 54.3%; Score 1952; DB 2; Length 631;
Best Local Similarity 64.7%; Pred. No. 1.5e-58;
Matches 410; Conservative 51; Mismatches 85; Indels 88; Gaps 3;
Qy 164 LRLISLEMLKRNKRETKRMGMMAKQEGMEKMLQVTSRLESQKIAQLEGKLVYSIEKE 223
Db 1 MRALSLEMLKRNKRETKRMGMMAKQEGMEKMLQVTSRLESQKIAQLEGKLVYSIEKE 60
Qy 224 KIDKSETEKLEYLEETISCASDQVEKYKLDIAQLEENLKEKNDEILSLKOSLEENIVIL 283
Db 61 KIDKCEETEKLEYLETISCASDQVEKCKVDIAQLEEDLKEKDRILSLKOSLEENIT-F 119
Qy 284 SKQVEDLVKCOLLEKEKEKEDHVNRRHNEHNLNADMQLKQKFILOEQEHKELQKELQI 343
Db 120 SKQIEDLVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALERQYEYKLOQKELQS 179
Qy 344 DSLQOKEKELSSSLHQKLCSPQEEVMVEKNLFEELKOTLDLDELKLOQKEQAEALYKOL 403
Db 180 QSLQOKEKELSLARLQOQLCSPQEEMTSEKNVFKELKLALAEALDAVQOKEQSERLVKOL 239
Qy 404 EEEAKSRAEELKLEELKLGKEAELEKSSAAHTQATLL----- 441
Db 240 EEERKSTAEQLTLDNLREKEVELEKHIAHAQAAILIAQEKYNDTAQSLRDVTAQLESV 299
Qy 442 ----- 441
Db 300 QEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQL 359
Qy 442 --LOEKYDSMVQSLDVTQAFESYKALTASEIEDLKLENSLQEKAAKAGNAEDVQHOI 499
Db 360 ESVQEKYNDTAQSLRDVTAQLESYKSSLTKEIEDLKLENLTLQEKVAMAESVEDVQOOI 419
Qy 500 LATESSNOEYVRMLDLQTKSALKETEIKETIVSFLOKITDLONLQKQEEDEFKQLEDE 559
Db 420 LTAESTNOEYARMVQDLOQRSTLKEEELKEITSSFLKXITDLKNLQROQDEDFRQLEBK 479
Qy 560 EGKRAEKENTTAELTEINKWRLLYELYNKTKPFQQLDAPFEVEKQALLNEHGAEOEL 619
Db 480 GKRYAEKENVMTLMEINKWRLLYDELYEKTQKPFQOQDAPFAEAEKQALLNEHGAETQEL 539
Qy 620 NKIRDSYAKLLGHONLKQKIKHVVKLKDENSQKSEVSKLRCOLAKKKQSETKLQEEELNK 679
Db 540 NKIRDSYAKLLGHONLKQKIKHVVKLKDENSQKSEVSKLRSQLVKKRQNELRLQGEJDK 599
Qy 680 VLGIKHFDPSPKAFHHESKENFALTKPLKEGNTNC 713
Db 600 ALGIRHFDPSKAFCHASKENF---TPLKEGNPNC 630
RESULT 3
S67593
transport protein USO1 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein D252; protein YDL058w
C:Species: Saccharomyces cerevisiae
C>Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 21-Jul-2000
C:Accession: S67593; A38455; S30782
R:Blöcker, H.; Brandt, P.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67587
A:Accession: S67593
A:Molecule type: DNA
A:Residues: 1-1790 <BLO>
A:CROSS-references: EMBL:Z74106; NID:g1431058; PID:e253003; PID:g1431059; MIPS:YDL058
A:Experimental source: strain S288C
R:Nakajima, H.; Hirata, A.; Ogawa, Y.; Yonehara, T.; Yoda, K.; Yamasaki, M.
J. Cell Biol. 113, 245-260, 1991
A>Title: A cytoskeleton-related gene, USO1, is required for intracellular protein tra
A:Reference number: A38455; MUID:91185402; PMID:2010462
A:Accession: A38455
A:Molecule type: DNA
A:Residues: 1-389,'TA',392-724,'S',726-1790 <NAK>
A:CROSS-references: GB:X54378; NID:g4777; PIDN:CAA38253.1; PID:g4778

Db 611 AADRVL--LETSVNEINLSQNESKEKVSQD--IQIAKTELLSAAEAQAQADLQ 666
Qy 413 -ELKLEKLGKAEAELEKSSAAHTQATLLQERYDSVMQVSLSDVTAQFESYKALTASEI 471
Db 667 NHDTAQNALQDKHQLNKITTDQVTAQKLDQKQEHCSQLESILKEYKYLSL-EQKT 725
Qy 472 EDL-----KLESSLQEKAAKAGNAEDVHQ-----ILATESSNQ-----EYVRML 513
Db 726 EELEGQIKLEADSLVSKASKE-QALQDQOQORQNTDLELRATLSQLQMEKEIVSS 784
Qy 514 -LDLQTSALKETELKEITVTFLOKTIYDLQNLQKQBEDFKQLEDEGRKAERKENTAE 572
Db 785 RLDLQKKS-----FALESIKOKLT-----KOEED--KQLTKQDFETLSQET----- 823
Qy 573 LTEEINKWRLLYEELYNKTRPFQLODAFEVEKQALLNEHGAQOELNKIRDSY----- 626
Db 824 -----KIOHEELNRIQTITVTELQKVMEKEALMTSLTVADKLSKVSDSLKNKSKSE 875
Qy 627 -----AKLLGHQNLKQKIKHVVKLDKENSOLKSEVSKLRQLAKKQKQSETKLOEEL 677
Db 876 FEKENQKGAAILDEKTKELKHQLOVQMENT-LK-EQKELKKSLEKEKEASHQLKLEL 933
Qy 678 NKV 680
Db 934 NSM 936
RESULT 7
S38173
myosin-like protein MLP1 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YKR095w; protein YKR415
C:Species: Saccharomyces cerevisiae
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 21-Jul-2000
C:Accession: S38173; S40647; S31207
R:Baladron, V.; Balastera, J.P.G.; Bou, G.; del Rey, F.; Esteban, P.F.; Garcia-Cantalejo, submitted to the Protein Sequence Database, March 1994
A:Reference number: S38158
A:Accession: S38173
A:Molecule type: DNA
A:Residues: 1-1875 <BAL>
A:Cross-references: EMBL:X73541; NID:g450550; PIDN:CAA51948.1; PID:g450554
A:Experimental source: strain S288C
R:Bou, G.; Esteban, P.F.; Baladron, V.; Gonzales, G.A.; Cantalejo, J.G.; Remacha, M.; Jh
Yeast 9, 1349-1354, 1993
A:Title: The complete sequence of a 15 820 bp segment of Saccharomyces cerevisiae chrom
A:Reference number: S40644; MUID:94205265; PMID:8154186
A:Accession: S40647
A:Molecule type: DNA
A:Residues: 1-1875 <BOU>
A:Cross-references: EMBL:X73541; NID:g450550; PIDN:CAA51948.1; PID:g450554
A:Experimental source: strain S288C
R:Koelling, R.; Nguyen, T.; Chen, E.Y.; Botstein, D.
Mol. Gen. Genet. 237, 359-369, 1993
A:Title: A new yeast gene with a myosin-like heptad repeat structure.
A:Reference number: S31207; MUID:93247549; PMID:8483450
A:Accession: S31207
A:Molecule type: DNA
A:Residues: 1-300, 'A', 302-1875 <KOE>
A:Cross-references: EMBL:L01992; NID:g171958; PIDN:AAA34783.1; PID:g171959
C:Genetics:
A:Gene: SGD:MLP1
A:Cross-references: SGD:S0001803; MIPS:YKR095w
A:Map position: 11R
Query Match 11.6%; Score 417; DB 2; Length 1875;
Best Local Similarity 24.1%; Pred. No. 4.9e-07;
Matches 190; Conservative 145; Mismatches 292; Indels 162; Gaps 30;
Qy 29 LEVLKGPVSFOKSRFOQKQESKQNLNVND----KDTILPASARKVKSESSEKES--QKND 82
Db 816 LSELKETS-QKDHHIKOLEED--NNSNIEWQNKIEALKDYEVITSVDSKQTDIEKILQ 873
Qy 83 KDLKILEKEI---RVLLQERCAQDRRIQD--LETELEKMEARLNAA-----LREKT 128

Db 874 YKVSLEKEIEEDKIRLHTYNVMDETINDSLRLEKSKINLTDAVSQIKEYKDYETT 933
Qy 129 SLSANNATLEKQLTELFTNELLKSKTSENGNQ-KNL-----RIISLELMKLRN 176
Db 934 SOS-----LOQTSKLDSEFXDFTNQIKNLTDEKTSLEDKISLKEQMFNLNN 981
Qy 177 KRETKMRGMMAKQEGMEMKLOVTRSEESQKIAQLSGKLVSTSEKEXIDKS----- 229
Db 982 ELDLQKGMKEKADFKKRISILQNNKNEVEAVKSEYKSLKIQND-LDOOTIYANTAO 1040
Qy 230 -----ETEKLLLEYIEEISCASDVKEYK-----LDIAQLEENKKNDELSLKQSLE 277
Db 1041 NNYEQELQKHADVSKTISELREQLHTYKQGVKTLNLSRDQLENALKENKSWSSOKESIL 1100
Qy 278 ENIVILSKQVEDLVNVCOLLEKEKEDHVRNREHNENLNA-----EMQNLAQK 325
Db 1101 EQLDLSNRIEDLSQNKLLYDQIYTAADKEVNNSTPGFLNNILITLRERDILDTK 1160
Qy 326 FILEQOEHEKLOQKELQIDSLLQO-----EKELSSSLHOKLCSFOEE-----MVK 370
Db 1161 VTVAERDAKMLRQKLSLMDVELQDARTKLDNSRVEKENHSSIIQOHHDI MEKLNQLNLIR 1220
Qy 371 EKNL-FEEL-----KQTLDELQKQKEEQAEERLYKQLEEEAKSRAELKLEEKL- 421
Db 1221 ESNITLRNLENNNNKKELQSELQKQNVAPIESELTKALKYSMOEKEQELKLAKEVH 1280
Qy 422 ---KGKAEAELEKSSAAHTQATLLQEKYDSVMQVSL-----DVTQAFESYKALT 467
Db 1281 RWKRRSDILEK----HQLSSSDYKLESEIENKELENKERQGAEEKFNLRQA 1336
Qy 468 ASETEDKLENSLSQLEKA-----AKAGNAEDVOHOILATFESSNQEVVRML 513
Db 1337 QERLTKSLQSDSTEQVNSLRDAKNVLENSANARIEELQNAKVAQGNQLERIKL 1396
Qy 514 L-DLQTSALKETELKEITVTFLOKTIYDLQNLQKQBEDFKQLEDEGRKAERKENTAE 572
Db 1397 QEDAEKASRELQAKLEESTTSVESTINGLNEEITTLKEIEKQROIQOOLQATSANEQND 1456
Qy 573 LTEEINKWRLLYEELYNKTRPFQLODAFEVEKQALLNEH-GAAEQELNKRIRDSVAKLIG 631
Db 1457 LSNIVESMKKSFEEL-DKIK-----FIKEQTQEVNKEILEAQERLQPSN-----IN 1501
Qy 632 HONLKQKIKHVVKLDKENSOLKSEVS-KLR-COLAKKQKQSETKQLEELNKVLGIKHFDP 689
Db 1502 MEEKKKWE-----SEHEQEVSQKIRAEAEALKRIRLPTTEKINKIIEERKEELE 1552
Qy 690 KAFTHESKE 698
Db 1553 KEFEKVEE 1561
RESULT 8
T14867
Interaptin - slime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 11-May-2000
C:Accession: T14867
R:Rivero, F.J.; Kuspa, A.; Brokamp, R.; Matzner, M.; Noegel, A.A.
J. Cell Biol. 142, 735-750, 1998
A:Title: Interaptin, an actin-binding protein of the alpha-actinin superfamily in Dic
ts.
A:Reference number: 218248; MUID:98365468; PMID:9700162
A:Accession: T14867
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1738 <RIV>
A:Cross-references: EMBL:AF057019; NID:g3549260; PID:g3549261; PIDN:AAC34582.1
C:Genetics:
A:Gene: abpp
A:Introns: 173/2; 1680/1
Query Match 11.6%; Score 416; DB 2; Length 1738;

Best Local Similarity 23.5%; Pred. No. 5e-07;
Matches 173; Conservative 157; Mismatches 276; Indels 128; Gaps 27;

QY 43 REKQKESQNLNVKDDTTLPASARKVSKSEKSKQNDKDLKILEKIRVLLQERGAQ 102
Db 789 QYKQOOLS-SNSNIDQ-----LSTIIELSELKEKELNDKLEKEKQLQLOQEFQOL 843
QY 103 DRRIQ-DLETELEKEMARLNAALREKTSLSANNATLEKOLIELTRTNELLKSKFSENGQ 161
Db 844 NEKNQKHQDQLELLEKQLKQLOQEVQDLQNETNQSIENQNLNOLINK-----ENLNE 896
QY 162 KNLRLLSLEMLKRLNRKRETKMGMAKQEGMEMKLOVQORSLEESQKTAQLEGLVSTIE 221
Db 897 K-----EQELLQNLQNOQIEKIQFDQOEFKQNSINTELVNEKNEKLIQLOQDYDQLK 951
QY 222 KE--KIDESETEKLELLEYIEELSCASQVQEKYKLDIALEENLKEKNDIELSKOSLEEN 279
Db 952 QQNRSDKEND-D-LIE-----KENQLKSTONELNOLIEKNESSHKEQQLKQOSIEND 1003
QY 280 IYILSKQVEDLNVKQLEK-----EKEDHYNRNRHNEHNEAEMONLKQKQFILEQ 330
Db 1004 LIEKENQIQOL--QSQNLNROQOQSNQISEKQDQNLQLEKQ--FDQEQQLKQOSIEND 1060
QY 331 --OEHEKLOQKELQDLSLQKEKELSSSHQKLCFSQFQEMVKEKNLFEELKQTTDELDK 388
Db 1061 LFEKENQIQOLQSQNLQNEQKQSQNLSEKQDQNLQLEK--NESQKQEQQLKQOSIEND- 1117
QY 389 LQKQEQOARLVKQLEEEAKSRAE-----ELKLEELKKGKEAELEKSSAAHTQATLL 442
Db 1118 LIEKENQIQOLQLEQKQLOQSEIVSDNDKILEKQKQSQDQLLNDEKQKQDQOL 1177
QY 443 QEK---YDSM-----VQSLVDVTAQFESYKALTAIEIEDLKLENSSLQEK 484
Db 1178 QDKQIEFDQQLTFNQFNKNDKDSQFIQDDQKQLOQSQ-----QDLNQLKQENQEKQ 1233
QY 485 AKAKNAEDVQHOILATESNQEVYVRLMLDLQTSALKETKEIKETVTSFKITDLO-- 542
Db 1234 LSEKDEKQSQFQENQEKQSEKDEKQSQSQNLQNLQNLNDEKQVQFSEKDEKQSQ 1293
QY 543 ---NOLKQDEDFRQDEGRKAERKENTTAETERINKRLWLYEELYNKTKPFQLOL 598
Db 1294 QDQLNQLKQENQEKQ-----SEKDEKQSQDQDQ-----LNDQIKRNEKLEK--- 1340
QY 599 DAFEVEKQALLNEHGAQQLNKRIRDSYAKLLGHQNLKQIKHYVVKLDKENSQKSE--V 656
Db 1341 ---EKEEQLLKIQDNDQSQSQQLQLEKLEKQENQQLQKQENQINQNLQOQSQSNEII 1397
QY 657 SKLQOLAKKKQSE-----TKQBEL-----NKVLGI-----KHFD 687
Db 1398 QOLKQDQLKQOQOQOQENNEKEIERLQIEBQLKQOQOQIDQSELSNKEIKIOTTQQEFD 1457
QY 688 PSKAFHESKENFALK 703
Db 1458 ---QLSHNRKQOLHLQ 1471

RESULT 9
A59252
myosin heavy chain, nonmuscle, form IIB - human
N:Alternate names: myosin type I0; NMHC-B
N:Contains: myosin ATPase (EC 3.6.4.1)
C:Species: Homo sapiens (man)
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-Apr-2002
C:Accession: A59252; B61231; G02055
R:Phillips, C.L.; Yamakawa, K.; Adelstein, R.S.
J. Muscle Res. Cell. Motil. 16, 379-389, 1995
A:Title: Cloning of the cDNA encoding human nonmuscle myosin heavy chain-B and analysis
A:Reference number: A59252; MUID:96025307; PMID:7499478
A:Accession: A59252
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-1976 <SIM>
A:Cross-references: GB:M69181; NID:g641957; PIDN:AAA99177.1; PID:g641958

A:Experimental source: clone lib Lambda Zap II adult human T-cell library; cell line
A:Note: between nucleotides 1942-1943 in mRNA encoding human brain MHC-B there is an
R:Simons, M.; Wang, M.; McBride, O.W.; Kawamoto, S.; Yamakawa, K.; Gdula, D.; Adelste
Circ. Res. 69, 530-539, 1991
A:Title: Human nonmuscle myosin heavy chains are encoded by two genes located on diff
A:Reference number: A61231; MUID:91316803; PMID:1860190
A:Accession: B61231
A:Molecule type: mRNA
A:Residues: 63-237, 'K', 239-664, 'L', 666-722 <SI2>
A:Cross-references: GB:M69181; NID:g641957
R:Weir, L.
submitted to the EMBL Data Library, August 1995
A:Reference number: H00753
A:Accession: G02055
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-81 <WEI>
A:Cross-references: EMBL:U34304; NID:g1143217; PIDN:AAA84880.1; PID:g1143218
C:Genetics:
A:Gene: GDB:MYH10
A:Cross-references: GDB:127350; GDB:G00-127-350; OMIM:160776
A:Map position: 17p13-17p13
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; nucleo
F:88-771/Domain: myosin motor domain homology <MMO>
F:178-185/Region: nucleotide-binding motif A (P-loop)
F:559-572/Region: actin binding #status predicted
F:633-647/Region: actin binding #status predicted
F:123/Modified site: N6,N6-trimethyllysine (Lys) #status predicted
F:184/Binding site: ATP (Lys) #status predicted
F:701,711/Active site: Cys #status predicted

Query Match 11.5%; Score 414; DB 2; Length 1976;
Best Local Similarity 23.4%; Pred. No. 6.5e-07;
Matches 199; Conservative 150; Mismatches 310; Indels 190; Gaps 29;

QY 32 LKGPVSFQKQKQKQKQNLNVKDDTTLPASARKV-----KSS 72
Db 1074 LQAQIDELKQLAKKKEELQALRGDDDETLHKNNALVVRVRELOAQIAELQEDFSEKAS 1133
QY 73 ESKRESQND--KDKLILEKEIRVLLQERGAQ-----DDR 105
Db 1134 RNKAEKQKRDSEELKALTELEDLTAAQOELTRKEQEVAKLKALBETKKNHEAQ 1193
QY 106 IQDLE---TELEKEARLNAALREKTSLSANNATLEKOLIELTRTNELLKSKFSENGQ 161
Db 1194 IQDMQRHATALEELSEQLQAKKFKANLEKNKQGLTDNKLCEVAVKQVKAESK 1253
QY 162 K-----NLRI--LSLEMLKRLNK-----RETKRMGMMAKQ--GM 192
Db 1254 RKKLDAAQVQELHAKVSEGDRLRVLEAKKSLQNLNDNVSTLLEAKKGIKFAKDAASL 1313
QY 193 EMKLOVQORSLEESQKTAQLEGLVSTIEKEKIDSEKTEKLELLEYIEISCASQOYEKK 252
Db 1314 ESQLODTELLQOETROKLNSSRIRQLEEK-----NSLQEQEKEEARKNLEKOV 1366
QY 253 LDIAQLEENLEKNDIELSKQSLKE-----NIVILSKOVEDLNVKCOLLEKEKEDHV 305
Db 1367 LAIQSQLADTKKKVDDDLGTTIESLEAKKLLKDAEALSQRLEKALAYDKLETK--- 1422
QY 306 NRNRHNEHNAEMONLKQKFILEQOHEKLEKQKELQDLSLQOQEKELSSSHQK----- 360
Db 1423 NQLQOELDDLAVDLDH-----QROVASNLEKQKQKFDQLLAEKKSISARYAEERDRAE 1475
QY 361 -----LCSFQEMVKEKNLFEELKQTTDELDKLOOKEQEARLVKQLEFEAK 408
Db 1476 AEAREKETKALSARALBEALKEAEKPEFRQNKQLRADMEDLMSSKDDYGVKNVHELEKSR 1535
QY 409 SRABELKLEELKKGKEAELEKSSAAHTQATLLLQEKYVDSMVQSLVEDVTAQFESYKALTA 468
Db 1536 ALEQQVEEMFTQLELEDELEQATEDAKLRLEVNQMAKQAFERDQTRDQNEEKRLLI 1595
QY 469 SEIEDL--KLENSSLQEKAAKAGK-----NAEDVOHQILATESNQEVYVRLMLDLQTKSA 521

Db 1596 KQVRELEAELEDERKQRALAVASKKMEIDLKLEAQTEAANKARDEVIKQRLKLAQMK 1655
Qy 522 LKETEIKETIYS---FLQ-----KTTDLQNLKQOE-----EDFRKQLED 558
Db 1656 DYQRELEARASRDEIFAQSKESKUKLSAEILQLOEEELASSERARRHAEQERDELAD 1715
Qy 559 E-----EGRKA---EK-----ENTTAELTEINKWRLLYEELYNKTKPFQQLDADFVEKQA 607
Db 1716 EITNSASKSALLDEKRLEARIAQLEEEEOSNMELLNDRFKTTLQVDTLNAE--- 1772
Qy 608 LLNHEGAAQEOQLKIRSYAKLLGHON--LKQIKHV-----VKLKDNSOLSKSEVSKLR 660
Db 1773 LAASERSAQAQS-----DNARQOLERONKELKAKIQLEAGVYSKFKATISALEAKIGOLE 1827
Qy 661 COL---AKKKOSETKLOEELKNVL-----GKHPDPSKAFHESKENF-ALKTPLK 707
Db 1828 EQLQOEAKERAANKLVRRTEKKLKEIFMQVEDERRHADQYKEQMEKANARKOLKROLE 1887
Qy 708 EGNNTNCYRA 716
Db 1888 EAEBEATRA 1896

RESULT 10
A26655
myosin heavy chain [similarity] - slime mold (Dictyostelium discoideum)
N:Contains: myosin Anpase (EC 3.6.4.1)
C:Species: Dictyostelium discoideum
C>Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 19-Apr-2002
R:Warrick, H.M.; De Lozanne, A.; Leinwand, L.A.; Spudich, J.A.
Proc. Natl. Acad. Sci. U.S.A. 83, 9433-9437, 1986
A:Title: Conserved protein domains in a myosin heavy chain gene from Dictyostelium discoideum
A:Reference number: A26655; MUID:87092266; PMID:3540939
A:Accession: A26655
A:Molecule type: DNA
A:Residues: 1-2116 <WAR>
A:CROSS-references: GB:M14628; GB:M11938; NID:g167834; PIDN:AAA33227.1; PID:g167835
R:DeLozanne, A.; Lewis, M.; Spudich, J.A.; Leinwand, L.A.
Proc. Natl. Acad. Sci. U.S.A. 82, 6807-6810, 1985
A:Reference number: A24728; MUID:86016788; PMID:3901008
A:Accession: A24728
A:Molecule type: mRNA
R:Wagle, G.; Noegel, A.; Scheel, J.; Gerisch, G.
FEBS Lett. 227, 71-75, 1988
A:Title: Phosphorylation of threonine residues on cloned fragments of the Dictyostelium
A:Reference number: S00250; MUID:88112226; PMID:2828113
A:Accession: S00250
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1734-1893 <WAG>
C:Comment: The rod domain is highly periodic, containing a pattern of 7-residue repeats
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: actin binding; ATP; coiled coil; hydrolase; nucleotide binding; P-loop; phosphatase
F:1-818/Domain: globular head <HEAD>
F:89-747/Domain: myosin motor domain homology <MMOT>
F:179-186/Region: nucleotide-binding motif A (P-loop)
F:819-2116/Domain: alpha-helical rod <ROD>

Query Match 11.5%; Score 414; DB 2; Length 2116;
Best Local Similarity 22.3%; Pred. No. 7e-07;
Matches 191; Conservative 178; Mismatches 275; Indels 214; Gaps 36;

Qy 8 LKRFNDPSGCAPSGAYD-VKTLVLPKVPVSFQK-----SOFKQOKESKONLVNDK--- 58
Db 959 MKRVND-----GQSDTISRLEKIKD--ELQKEVEELTESFSSESXDKGVLEKTRVRL 1008
Qy 59 -----DTTLPASARKVSKESKESKNDKDLKILEKEIRVLVLPQRGADRRQIDLETLEL 113
Db 1009 QSELDLTVRLDSETKOKSELLRQKKLEELKQVQEEALAAETAAKLAQEAANKKLOGEY 1068

Qy 114 EKMEARLNAALREKRTISLSANNATLEKQILERTNELLKSFSENGNKNRIILSLELMK 173
Db 1069 TELNEKFNSEVTARSNVEKSKTTLESQLVAV--NNELDEEKNRDALKKKKALDAMLEE 1126
Qy 174 LRNRETKMRGMMAQOEGEMK-----LQVTORS--LEESQGIQAQLEGKLVISIEKEKIDEX 228
Db 1127 MKDQLES-----TGGEKKSLYDLKVKQESDMELARNQISELQSTIAKLEKTKSTLE 1177
Qy 229 SETKLLLEYIEEISCASQOYEK-----LDI-----AQLEENLKEKNDEILSKQSLEENI 280
Db 1178 GEVARLQGELEAEQALAKSNVEKQKKKVELDLEDLSAQALAEATAAK-QALDKLKKLEBEL 1236
Qy 281 VILSKQVEDLVKVCQLEKEKEDHVNRRNHNENINAEOMNLKQKFIIEQQEHEKLOOKE 340
Db 1237 SEVQTLSEANNK-----NVNSD--STNKHLETFSNNLKLLEAEQAKAQALEKKR 1285
Qy 341 LOIDS-----LLOQEKELSSHLKQCSQOEMVKNLFEELBKQTLDELDKLOOKE 393
Db 1286 LGLESELKHVNEQLEEEKKQKESNEKRVKDLKEVESELDQIEEVSASKAVTEAKNKKE 1345
Qy 394 BOAERLVKOLEEAKSR---AEELKLLKEK-----LKGKEAELEKSSAAHTQATLLQ 443
Db 1346 SELDEIKROYADVSSRDKSVQQLTKLQAKNEELNATAEAEQGLDRAERSKKKAEFDLE 1405
Qy 444 EKYDSMVQSLDVTQ-QFESYKAL-----TASEIEDLK----- 475
Db 1406 E---AVKNLEETAKKVKAEKAMKKAETDYRSTKSELDDAKNVSSQYVOIKRLNEELS 1461
Qy 476 -----LE-----NSSLOEK-----AAKAGKNAEDVQHOILATES 504
Db 1462 ELRSVLEADRCNSAIKAKKTAESALESKDEIDAANNAKAKERKSKELEVRVALEEE 1521
Qy 505 S-----NOEYV-----RMLDLQTKSALKETEIKE----- 529
Db 1522 SLEKSGTVNVFEIRKKDAEIDDLRARLDRETESRIKSDDEKKNTKQKOFADLEAKVEAQ 1581
Qy 530 ---ITVSFLQK-----ITDLQNLKQOE-----EDFRKQLEDE--EGRKAKEKNTTABEL 574
Db 1582 REVVTIDRLKKLESIIIDLSTQDLDTETKSRIKIEKSKKLEQTLAERRAAEGSSKAAD 1641
Qy 575 EEINKWRLLYEELYNKTKPFQQLDADFVEKQALLNEHGAQEOQLNKIRDSVAKLLGHON 634
Db 1642 EEIRK-----QVMQOEVDELRAQLDS-----EAAALN-----ASEK--KIKSLVAEY---DE 1682
Qy 635 LKQIKTHVVVLLKDE---NSOLKSEVSKRLQAKQKQSETKLOEELNKLVLGKIKHFDPSK 690
Db 1683 VKEQLEDEILAKDLKLVKAKRALEVELEBVRQLOEEEDSRSSEL-EDSKRRLTTEVEDIKK 1741
Qy 691 AFHHEKSNFALKTPLKE 708
Db 1742 KYDAEVEQNTKLDEAKKK 1759

RESULT 11
A33977
myosin heavy chain, nonmuscle - chicken
N:Contains: myosin Anpase (EC 3.6.4.1)
C:Species: Gallus gallus (Chicken)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 19-Apr-2002
C:Accession: A33977; S06116; A43422
R:Shohet, R.V.; Conti, M.A.; Kawamoto, S.; Preston, Y.A.; Brill, D.A.; Adelstein, R.S.
Proc. Natl. Acad. Sci. U.S.A. 86, 7726-7730, 1989
A:Title: Cloning of the cDNA encoding the myosin heavy chain of a vertebrate cellular
A:Reference number: A33977; MUID:90046668; PMID:2813355
A:Accession: A33977
A:Molecule type: mRNA
A:Residues: 1-1959 <SHO>
A:CROSS-references: GB:M26510; NID:g212382; PIDN:AAA48974.1; PID:g212383
R:Katsuragawa, Y.; Yanagisawa, M.; Inoue, A.; Masaki, T.
Eur. J. Biochem. 184, 611-616, 1989
A:Title: Two distinct nonmuscle myosin-heavy-chain mRNAs are differentially expressed
s.
A:Reference number: S06116; MUID:90032648; PMID:2806244

A:Accession: S06116
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 716-1008 <KAT>
A:Cross-references: GB:X17589
A>Note: this translation is not annotated in GenBank entry GGMCFMHA, release 114
R:Hodge, T.P.; Cross, R.; Kendrick-Jones, J.
J. Cell Biol. 118, 1085-1095, 1992
A:Title: Role of the COOH-terminal nonhelical tailpiece in the assembly of a vertebrate
A:Reference number: A43422; MUID:92381096; PMID:1512291
A:Accession: A43422
A>Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1900-1959 <HOD>
A:Experimental source: brush border
A>Note: sequence extracted from NCBI backbone (NCBIP:111947)
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; nucleotide
F:84-764/Domain: myosin motor domain homology <MMOT>
F:174-181/Region: nucleotide-binding motif A (P-loop)
F:552-565/Region: actin binding #status predicted
F:626-640/Region: actin binding #status predicted
F:837-1936/Domain: coiled coil #status predicted <COI>
F:837-1277/Region: S2
F:1278-1959/Region: light meromyosin
F:1937-1959/Domain: carboxyl-terminal <CBT>
F:125/Modified site: N6,N6-trimethyllysine (Lys) #status predicted
F:180/Binding site: ATP (Lys) #status predicted
F:694,704/Active site: Cys #status predicted

Query Match 11.4%; Score 409; DB 1; Length 1959;
Best Local Similarity 23.0%; Pred. No. 9.5e-07;
Matches 189; Conservative 159; Mismatches 289; Indels 184; Gaps 32;

Qy 37 SFQSQRFQKQESQNLNVKDTLPASARKVSKSEKESQK--NKKDLKILEKEIRV 94
Db 873 TFO-AQLMAEKQLOQLQAEALCAEAETRLAKKQLEETCDLEARVEEERC 931

Qy 95 --LLQGAQDRRIODELEKME-ARLNAAAREKTS-----LSANNATLEKOLIET 145
Db 932 QHLQAEKKMQNQIOLEOLEEESARQKLOLEKVTYEAFLKLEEDVIVLEDONLKL 991

Qy 146 RTNELLSKFSF-NGNQNLRLSLLEMLKLRNRE---TKRMGMMAKQEGMEMKLVQTV 201
Db 992 KEKKLLDRMSEFTNLTEEEKSKSLAKKNHEAMITDLERLRREKQKOLEKTRR 1051

Qy 202 SLE-----ESQKIA-----QLEGLVSTKE-----KIDE-KS 229
Db 1052 KLEGSSDLHDOIAELQIAELKIQLSKKEELOALARVEEAAQKNMALKKIRELS 1111

Qy 230 EPEKLEYTEITSASDOVEKYKLDIAOLEENLEKNDEIL-----SLKQSLNENVI 282
Db 1112 QITELOEDLESASRNKAQKRDGLGEELEALTELEDLDSTAAQELSKRQEVTV 1171

Qy 283 LSKQVED-----LNVKCOLLEKEKEDHVN 306
Db 1172 LKKTLEDAKTHAQIQEMRQKHSQAIBELABQLQTKRVKRVANLEKAKQALESERAE 1231

Qy 307 R-----NREH-NEMLNEMQNLKQFI-----LEQOEHEKLOOKELOIDS---LLQQ 349
Db 1232 EVKVLQKGDGAEHRKKRVDAQLQELQKFTGERVKVTKLEAERVNKQLVDNVTGLNQ 1291

Qy 350 EKELS-----SSLRQLCSFQEMVKEKNL---FEPELQTLDELKLOOKEQAE 399
Db 1292 SDSKSIKLAKDFALESQLODQELQOETRLKLSFSTYKQTEDEKNALEOLEE 1351

Qy 400 VKQLEE-----EAKSRAE-----LKLEELKKEAELEKSSAAHTQ---ATLL 442
Db 1352 KRNLEKQISVLQQAQVAREARKMDGGLGCLTAEEAKKQLQKLDLESLTQRYE 1411

Qy 443 QEKYDSMQSLDVTVAQFESYKALTASIE-----DLKLENSSLOEKAA---KAG 491
Db 1412 EKTTRLOQELDDTAVDLD-HQRTVSNLEKKQKKFDGLAEENKISAKYAEERDRAE 1470

Qy 492 AEDVQHILATESNQEYVRMLLDLQTSALKETIKETVS-----FLQKITDLOQL 545
Db 1471 AREKETKALSARALEEALEQKAELEVRNQKQRTMEDLMSSKDDVGKSVHELEKRAL 1530

Qy 546 KOQEDFRKQLEDEGRKAENKTAETETINKWR-----LLYEELYNKTPFQI- 596
Db 1531 EQQVEEMKTOLEDELOQATED--AKLRLEVNQQAQAFDRDLGLRDEQNEERKOLI 1588

Qy 597 -OLDAFEVEKQALLNHEGNAQPOLNKRDSYAKLGH-----QNLKQKIKHVVKLKDENS 650
Db 1589 ROVREMEVELEDERKQSIARVAARKKLELDKDLKLESHDITANKNDEAIKHVRKLAQAK 1648

Qy 651 QLKSVSKLRQ----LAKKKQSETK-----LQEEEL 677
Db 1649 DYMRELEDTYSREILAQAKENKKLKSMEAEMLQLEEL 1689

RESULT 12
B43402
myosin heavy chain-B, neuronal - chicken
N:Contains: myosin ATPase (EC 3.6.4.1)
C:Species: Gallus gallus (Chicken)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 19-Apr-2002
C:Accession: B43402; A43402
R:Takahashi, M.; Kawamoto, S.; Adelstein, R.S.
J. Biol. Chem. 267, 17864-17871, 1992
A:Title: Evidence for inserted sequences in the head region of nonmuscle myosin spec
yosin.
A:Reference number: A43402; MUID:92388144; PMID:1355479
A:Accession: B43402
A:Molecule type: mRNA
A:Residues: 1-2007 <TAK>
A:Cross-references: GB:M93676; NID:g212448; PIDN:AAA48988.1; PID:g212452
A>Note: the sequence of residues 212-221 and 632-652 and the corresponding nucleotide
A:Accession: A43402
A:Molecule type: mRNA
A:Residues: 1-211,222-631,653-2007 <TA2>
A:Cross-references: GB:M93676; NID:g212448; PIDN:AAA48985.1; PID:g212449
A>Note: sequence extracted from NCBI backbone (NCBIN:112864)
C:Comment: Alternatively spliced segments 1 and 2 are found exclusively in nonmuscle
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: actin binding; alternative splicing; ATP; coiled coil; hydrolase; methyl
F:1-2007/Product: myosin heavy chain-B, neuronal #status predicted <MYS>
F:1-211,222-631,653-2007/Product: myosin heavy chain-B, nonmuscle #status predicted <
F:88-802/Domain: myosin motor domain homology <MMOT>
F:178-185/Region: nucleotide-binding motif A (P-loop)
F:212-221/Region: alternatively spliced segment 1 #status experimental
F:559-593/Region: actin binding #status predicted
F:632-652/Region: alternatively spliced segment 2 #status experimental
F:692-714/Region: actin binding #status predicted
F:875-2007/Domain: coiled coil #status predicted
F:875-1315/Region: S2
F:1316-2007/Region: light meromyosin
F:129/Modified site: N6,N6-trimethyllysine (Lys) #status predicted
F:184/Binding site: ATP (Lys) #status predicted
F:732,742/Active site: Cys #status predicted
F:1954/Binding site: phosphate (Thr) (covalent) #status predicted
F:1987/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 11.4%; Score 409; DB 1; Length 2007;
Best Local Similarity 23.6%; Pred. No. 9.7e-07;
Matches 199; Conservative 152; Mismatches 314; Indels 178; Gaps 30;

Qy 32 LKGPVFSQKSFQKQESQNL-NVDDTTLTPASARKV-----KSS 72
Db 1105 LOAQIEELKTLQAKKEELQALARGDEEAVQKNNALKVIRELQQAIAELEDSEKAS 1164
Qy 73 ESKESQKND--KDLKILEKIRVLLQERGAQ-----DRR 105
Db 1165 RNKAQKQKRDLSLEALKTELEDLTDTAAQQLRTKREQVEALKAIBETKNHEAQ 1224
Qy 106 IQDLE----TELEKNEARLNAAAREKTSLSANNATLEKQLIELTRTNELLSKFSQNG 161

1225 IQEIRQRHATALELSQEQAKRFKANLEKNKQGLSDNKNELACEVQKVQVQVKAESHK 1284
QY 162 K-----NLRI--LSLEMLKLRN-----RETMRGMMAQOE--GM 192
Db 1285 RKKLDAQVQELTAKVTEGERLVELABKANKLQNELDNVSSLLEAEAKKGIKFAKDAASL 1344
QY 193 EMKLQVTRQSRLEESQGIQALQEGKLVEIEKEIDSEKTEKLLYIEEISCASDOVBKYK 252
Db 1345 ESQLODTQELLQETROKLNLSRIROLEEK-----NNLQEQEPEEAKNLEKQM 1397
QY 253 LDI--AQLEENLKEKNDILSKQSLERNIVLSKQVEDLVNKKCOLLEKEKEDHVNRRNH 311
Db 1398 LALQAQAE--AKKVDVDDGLTIEGLEENKKLLKDMESLS---QRLB--EKAMAYDKLEKT 1452
QY 312 NENLNAEMQNLKQFLEQEQHEKQKQELQIDSLQEQELSSLLHOK----- 360
Db 1453 KNLQEQELDDLVMDLHQRIVSNLEKQKQKQFOOMLAEENKISARYAEERDRAEAEAREK 1512
QY 361 -----LCSFQEMVKEKNLFEELKQTLDELKQKQEAERLVKQLEEEAKSRAEEL 414
Db 1513 ETKALSARALEAEAEKFEERQNKQLRADMEDLMSKDDVGKNVHELEKSKRTLEQOV 1572
QY 415 KLEEKILKGKAELEKSSAAHTQATLLQEKYDSMWQSLVEDVTAQFYSYKALTASEIEDL 474
Db 1573 EEMRTQLELEDELQATEDAKRLLEVNNQAMKAQFERDLQARDQNEEKKRMLVKQVREL 1632
QY 475 --KLENSLSQEKAAKAK-----NAEDVQHOILATESNQYEVVMDLDTQKSKALTEI 527
Db 1633 EAELEDERKORALAAKAKKMEMDMLKQLEGIEAANKARDEAIKQRLKQAKMDYQREL 1692
QY 528 KEITVS---FLQ-----KITDQNLQKQOEEDF-----RKQLEDE----- 559
Db 1693 EEARSRDEIFAQSEKSEKLLKGLAEILQLOEFAERASERARRHAEQERDELADEIANS 1752
QY 560 EGRKA---EK---ENTTAETEEINKWRLLYEELYNKTKPPQLQDLDAFEVEKQALLNEHG 613
Db 1753 SGKSALLDEKRLREARIAQLEEELEEQSNMELNFRKTTLQVDTLNSF---LAGERS 1809
QY 614 AAQEQNL-----KIRDSYAKLGHQNLKQIKHVVKLKDENSQKSEVSKRLCOL--- 663
Db 1810 AAQSEANARQOLRQNLKAKL---QELGSEVKS---KFKATISTLEAKIAQLEEQLEQE 1864
QY 664 AKKQSETKLOEELNKVL-----GIKHFDPSKAFHHESKNF--ALKTPPLKEGNTNC 713
Db 1865 AKERAAANKLVRRTTEKLVKQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 1924
QY 714 YRA 716
Db 1925 TRA 1927

RESULT 13

A36014
myosin heavy chain, nonmuscle - fruit fly (Drosophila melanogaster)
N:Contains: myosin ATPase (EC 3.6.4.1)
C:Species: Drosophila melanogaster
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 19-Apr-2002
C:Accession: A36014; B36014
R:Ketchum, A.S.; Stewart, C.T.; Stewart, M.; Kiehart, D.P.
Proc. Natl. Acad. Sci. U.S.A. 87, 6316-6320, 1990
A:Title: Complete sequence of the Drosophila nonmuscle myosin heavy-chain transcript: cd
A:Reference number: A36014; MUID:90349606; PMID:2117279
A:Accession: A36014
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-2017 <KET>
A:Cross-references: GB:M35012
C:Genetics:
A:Gene: FlyBase:zip
A:Cross-references: FlyBase:FBgn0005634
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: actin binding; alternative splicing; ATP; coiled coil; hydrolase; methylated

F;1-2017/Product: myosin heavy chain, form I #status predicted <MAI>
F;46-2017/Product: myosin heavy chain, form II #status predicted <MA2>
F;135-815/Domain: myosin motor domain homology <MMOT>
F;225-232/Region: nucleotide-binding motif A (P-loop)
F;598-631/Region: actin binding #status predicted
F;705-727/Region: actin binding #status predicted
F;888-2017/Domain: coiled coil #status predicted <COI>
F;888-1328/Region: S2
F;1329-2017/Region: light meromyosin
F;176/Modified site: N6,N6-trimethyllysine (Lys) #status predicted
F;231/Binding site: ATP (Lys) #status predicted
F;745,755/Active site: Cys #status predicted

Query Match 11.3% Score 408; DB 1; Length 2017;
Best Local Similarity 24.7% Pred. No. 1.1e-06;
Matches 209; Conservative 134; Mismatches 286; Indels 218; Gaps 35;

QY 29 LEVLGPVSP--QKSQRFQKQKESQNLN-----VDKDT-----TLPA 64
Db 889 LEVTKQEEKLVQKEDELKQVREKLDTLAKNTQEVKRYQVALVEKTTLAEQLQAEIELCA 948
QY 65 SARKVKS-----SESKKESQKNDKDKILEKEIRVLL--QERGAQDRRIQDLETELEKMEA 118
Db 949 EAEESRRLMARKOELEDMQOELETRIEEEERVYALGGERKKLELNIQDLEQLEEEA 1008
QY 119 RLNAALREKTSLSA-----NNATL--EKOLIELTRTNELLSKFSSENGNQK 162
Db 1009 ARQKQLEKQVOLDKIKKYEDLALTDQNOQKLLKEKKLLE--ERANDLSQTLAEIEEKAK 1067
QY 163 NLRILSLEMLKLRNRE---TKMRGMMAQEGMEMKIQVTORSLE-----ESQG 208
Db 1068 H-----LAKLAKKHEATIELEERLHKDQQRQESDRSKRTIETEVADLKEQLNERRV 1120
QY 209 KIAQLEKLVSIKIEK-----KIDEKSETE-----KLLEYIEEISCASDQ 247
Db 1121 QVDEMQALAKREELTQTLRIDDESATKATAQAKQARELSQALAEQIEDLEAEKARAK 1180
QY 248 VEKYLDTAOLEENLK-----EKNDILSKOSLEENTV----- 281
Db 1181 AEKYRDLSEELAKNELLDLSLDTAAQQLRSKRQELATLAKSLEETVNHGVLAD 1240
QY 282 -----ILSKQVEDLVNKKCOLLEKEKED-----HVNRRNHEMNENLN-- 316
Db 1241 MRKHQSQELNSINDQNLNLRKAKTVLEKAKCTLEAENADLATELSRVNSSRQENDRRKQ 1300
QY 317 -----AEMQNLKQFLEQEQHEKQKQELQIDSLQEQELSSLLHOKLCSF 364
Db 1301 AESQIAELQVLAETIERARSEL---OEKCTKLOQEAENITNOL--EAEELKASAAVKSASN 1356
QY 365 QEEMVKE--KNLFEELKQTLDELKQKQEAERLVKQLEEE--EAKSRAEELKLE--- 418
Db 1357 MESQLTAAQQLLEETQKQLGLSSKLRQISEKALQEQLEEDDEAK--RNYERKLAEVTT 1415
QY 419 --EKLGKGAELKSSAAHTQATLLQEKYDSMWQSLVEDVTAQ---FESYKALTASEIED 473
Db 1416 QMQEIKKAEEDADLAKLEBGGKRLNKDIEALERQVKEKLLAQNDRDLDSKKKIQSELED 1475
QY 474 KLENSLSQEKAAKAGNAEDVQHOILATESNQYEVVMDLDTQKSKALTEIETIVS 533
Db 1476 ATIELEAQRTKVLELEKKQKNFD--KILAEKAKISEQIAQERDRTAEREA--REKETKVLVS 1533
QY 534 -----FLQKITDQNLQKQOEEDFRKQLEDEEGRKAEKENTTAELTEINKWRLLYEELY 588
Db 1534 RELDEAFDKIEDLENK-----RKTLQNELDLDLANTQGTADKNVHELEKAKRALESQ 1585
QY 589 NKTK-----PPQLQIDA---FEVEKQALLNEHGAQEQNLKIRDSYAKLGHQNLKQ 637
Db 1586 AELKAQNEELEDLQLTEDAKLRLEVNNQALRSQFE-----ROLLAK---EEGAE 1633
QY 638 KIKHVVK--LKDENSQKSEVSKRLCOLAKKQSKETKQLEELNKVGLGKHFDPSKAFHHES 696
Db 1634 KRRGLVQRLDLELDEERKQRTAAVASKKKLEGDLEK-----IETTMEHMKV 1683

[illegible]

Qy 289 ----DLNVKCOLLEKEKEDHVR-----NREH-NENLNAEQMLKOKFI-----LEQ 330
Db 1210 RLKGNLEKAKALGSEGNELANVKTLLQCKGDESEHKRKKVQAQLEQVYVTEGDRVS 1269
Qy 331 QEHEKLOQKEIDID---SLQOEKELSSLHQKLCSPQOEEMVKEKNLFEELKOTLDELQ 387
Db 1270 ELSEKANRLQVELDNVNSLLSQSDSKSIKLGKDFSTLESQFQDAQELLQEBETROKLSFST 1329
Qy 388 KLOQKEQAERLVKOLEEAKSR-----AEEL-KL 416
Db 1330 KLQMEDEKNGLLQLEEEEAENKLCQISTQSQMTDMKKMDENVGSLVTEELKK 1389
Qy 417 LEELKLGKEAELEKSSAAHTQ----ATLLQEKYD-----SMVQSLEDVTAQFESY 463
Db 1390 LQKLEAVNQFFEEKAAYDKLEKTKTFLQELDDISVDLDHQRQIYVSNLEKKQKKFDQL 1449
Qy 464 ----KALTA-----SEIEDL 474
Db 1450 LGEKAIKSAKYADRRDRSEAEAREKETKALALARALEEALEAKAELERLNKQLRTEMEDL 1509
Qy 475 -----KLENS--SLOEKAAGNAEDVQHOILATESNQYVYRMLDLQTKS 520
Db 1510 VSSKDDVGVSVHELEKSKRALEQQAEMKTOLEEELELQATEDAK---LRLEYNLQAMK 1566
Qy 521 ALKETEIKEITVSLQKITDILQNLKQOE---EDFRKQ-----LEDEEGR- 562
Db 1567 AQFERDLQGRDEQSEDKKQLVRQVKEWAELEDERKQRSIYVAARKKLEMDLADLEGQI 1626
Qy 563 ---KAEKENTAELE-----EINKWRLLYEELYNKTKPQOLDAFEVEKQALL 609
Db 1627 DSSNKNREDAIKQLKLAQIKDYQRELEDTRASRDDILSQSEKSEKKLSMEAEIMHQ 1686
Qy 610 NEHCA-----AQEQLNKIRDSYAKLLGHQNLKQIKHVVKLKDENSOLKSEVSKLR-- 660
Db 1687 EELAAERGRKQAQOERDELADEIANSKGALALEEKR--RLESRTAQLEEELEEEQGN 1744
Qy 661 COLAKKKQSETKLQ-EELNKVLGKIKHFDPSK---AFHHEKSENFALTKPLKE 708
Db 1745 TELYNDRLKXSTLQIDQINTDLTAERSNAQKNENARQOQMDRONKELTKLOE 1796

Search completed: December 20, 2002, 15:12:38
Job time : 38.7824 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:06:33 ; Search time 37.4263 Seconds
(without alignments)
2581.255 Million cell updates/sec

Title: US-09-685-010-47

Perfect score: 3598

Sequence: 1 MSPPKAPLRFNDPSCAPS.....LKEGNTNCRAPMECOESWK 725

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3598	100.0	725	AAW39165	Human RHAMM protei
2	3598	100.0	725	ABG60842	Human receptor for
3	3581	99.5	725	AAU11436	Human hyaluronica
4	2670.5	74.2	713	AAU11438	Rat hyaluronica aci
5	2623	72.9	794	AAU11437	Mouse hyaluronica a
6	1953.5	54.3	630	AAW39166	Mouse RHAMM protei
7	1953	54.3	631	ABG60843	Mouse receptor for
8	1952	54.3	631	AAW39167	RHAMM 1-2a isoform
9	1819.5	50.6	606	AAW39168	Receptor for hyalu
10	1694	47.1	351	AAW01052	Human umbilical ve

11	1339	37.2	476	14	AAW39165	Hyaluronan recepto
12	429.5	11.9	1411	17	AAW02258	Nucleolar/endosoma
13	414	11.5	1988	22	AAW40999	Human polypeptide
14	414	11.5	1988	22	AAW41000	Human polypeptide
15	413	11.5	1388	23	AAU79590	Human kinesin moto
16	413	11.5	1388	23	AAE14400	Human kinesin supe
17	409.5	11.4	1948	22	ABG21233	Novel human diago
18	407	11.3	2056	22	ABG59344	Drosophila melanog
19	406.5	11.3	1690	22	ABG61144	Drosophila melanog
20	406.5	11.3	1690	22	ABG61173	Drosophila melanog
21	404.5	11.2	1960	22	AAW78854	Human protein SEQ
22	404.5	11.2	2143	22	ABG01716	Novel human diago
23	398.5	11.1	1857	23	AAU84350	Protein MYH11 diff
24	398	11.1	1972	17	AAW00024	Smooth muscle myos
25	397	11.0	1963	22	AAW79838	Human protein SEQ
26	396.5	11.0	2117	22	AAU32040	Novel human secret
27	396.5	11.0	2192	18	AAW21732	LexA/NumA fusion p
28	396.5	11.0	2272	18	AAW21731	GAL4/HA/NumA fusio
29	395.5	11.0	2101	15	AAW47173	Sequence of the in
30	395.5	11.0	2101	22	AAW65799	Nuclear mitotic ap
31	395.5	11.0	2207	22	AAU32041	Novel human secret
32	394.5	11.0	1427	12	AAW10534	Human 160kD mediat
33	393.5	10.9	2115	21	AAW49937	Human NumA protein
34	393	10.9	795	23	ABW77430	Human tumour marke
35	393	10.9	976	22	AAW66581	Human SCP-1 muten
36	392	10.9	1392	20	AAW06999	Restin protein seq
37	391.5	10.9	2246	22	ABG05850	Novel human diago
38	391	10.9	1489	22	ABW59948	Drosophila melanog
39	390.5	10.9	1312	18	AAW22775	Human RAD50. Homo
40	390.5	10.9	1312	19	AAW71295	Human homologue of
41	389.5	10.8	2101	21	AAW49936	Human NumA protein
42	386	10.7	1939	23	ABW77096	Human alpha-myosin
43	381.5	10.6	1879	22	AAW25750	Human protein sequ
44	381.5	10.6	2482	16	AAW72826	Human mitotin. Ho
45	381.5	10.6	2482	19	AAW23996	Human mitotin amin

ALIGNMENTS

RESULT 1

AAW39165

ID AAW39165 standard; Protein: 725 AA.

XX AC AAW39165;

XX DT 27-APR-1998 (first entry)

XX DE Human RHAMM protein.

XX KW Hyaluronan receptor; receptor for hyaluronan acid mediated motility;
XX KW RHAMM; glycosaminoglycan; binding domain; human; oncogene; treatment;
XX KW growth factor; cell locomotion disorder; dementia; detection;
XX KW inflammatory disorder; autoimmune disease; diagnosis; prognosis.

XX OS Homo sapiens.

XX PH Key Location/Qualifiers

XX FT Region 478..480

XX FT Region /note= "Region indicated in specification"

XX FT Region 568..570

XX FT Region /note= "Region indicated in specification"

XX FT Region 589..591

XX FT Region /note= "Region indicated in specification"

XX FT Binding-site 636..646

XX FT Binding-site /note= "Hyaluronic acid binding domain"

XX FT Binding-site 658..667

XX FT Binding-site /note= "Hyaluronic binding domain"

XX PN WO9738098-A1.

XX PD 16-OCT-1997.

XX

PF 10-APR-1997; 97WO-CA00240.
 XX
 PR 10-APR-1996; 96GB-0007441.
 XX
 PA (MANI-) MANITOBA CANCER TREATMENT & RES FOUND.
 PA (UYMA-) UNIV MANITOBA.
 XX
 PI Entwistle J, Turley EA;
 DR WPI; 1997-512715/47.
 DR N-PSDB; AAV02800.
 XX
 XX Isolated human receptor for hyaluronic acid mediated motility - used
 PT to develop products for treating e.g. tumours, inflammatory
 PT disorders, dementia, AIDS, diabetes and auto-immune diseases
 XX
 PS Claim 16; Page 46; 66pp; English.
 XX
 CC This sequence represents the human hyaluronan receptor which is also
 CC known as the receptor for hyaluronic acid mediated motility (RHAMM).
 CC Hyaluronan is a large glycosaminoglycan that is ubiquitous in the
 CC extracellular matrix and whose synthesis has been linked to cell
 CC migration, growth and transformation. It interacts with cell surfaces via
 CC specific protein receptors, e.g. RHAMM, that mediate many biological
 CC effects. The RHAMM/hyaluronic acid interaction is involved in
 CC oncogene- and growth factor-mediated cell locomotion. The products can be
 CC used in the treatment of disorders involving cell locomotion, e.g. tumour
 CC invasion, birth defects, acute and chronic inflammatory disorders,
 CC Alzheimer's and other forms of dementia, including Parkinson's and
 CC Huntington's diseases, AIDS, diabetes, autoimmune diseases, corneal
 CC dysplasia and hypertrophies, burns, surgical incisions and adhesions,
 CC strokes and multiple sclerosis. They can also be used in e.g. CNS and
 CC spinal cord regeneration, contraception and in vitro fertilisation and
 CC embryo development. The products can also be used in detection, diagnosis
 CC and prognosis.
 XX
 SQ Sequence 725 AA;
 Query Match 100.0%; Score 3598; DB 18; Length 725;
 Best Local Similarity 100.0%; Pred. No. 5,3e-185;
 Matches 725; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSFPAKPLKRFNDPSCAPSPGAYDVKTLEVLKGPVSPQSKQRFPQKESKONLVNDKDT 60
 DB 1 MSFPAKPLKRFNDPSCAPSPGAYDVKTLEVLKGPVSPQSKQRFPQKESKONLVNDKDT 60
 QY 61 TLPSARKVKSSKESKOKNDKDLKLEKEIRVLQERGAQDRRIQDLELEKWEARL 120
 DB 61 TLPSARKVKSSKESKOKNDKDLKLEKEIRVLQERGAQDRRIQDLELEKWEARL 120
 QY 121 NAALREKTSLSANNATLEKQLTELRTNELLSKSFSENGNKNRLISLLEMLKRNKRET 180
 DB 121 NAALREKTSLSANNATLEKQLTELRTNELLSKSFSENGNKNRLISLLEMLKRNKRET 180
 QY 181 KMRGMMAKQEGMEMQLVQTSRLESQKIAOLEGKLVSIIEKEIDKSETEKLEYIEE 240
 DB 181 KMRGMMAKQEGMEMQLVQTSRLESQKIAOLEGKLVSIIEKEIDKSETEKLEYIEE 240
 QY 241 ISCASDQVEKYKLDIAOLEENLKENDELTLKQSLLENIVILSKQVEDLVNKCALLEKE 300
 DB 241 ISCASDQVEKYKLDIAOLEENLKENDELTLKQSLLENIVILSKQVEDLVNKCALLEKE 300
 QY 301 KEDVHNRNHNENLNAEQNLKQFLEQOEHEKLOQKELQIDSLLOQKELSSSLHQK 360
 DB 301 KEDVHNRNHNENLNAEQNLKQFLEQOEHEKLOQKELQIDSLLOQKELSSSLHQK 360
 QY 361 LCSFOEMVYKKNLFEELKQTLDELKLOQKEQAEERLVQLEEEAKRAEELKLESEK 420
 DB 361 LCSFOEMVYKKNLFEELKQTLDELKLOQKEQAEERLVQLEEEAKRAEELKLESEK 420
 QY 421 LKGRKAELEKSSAAHTQATLLLOEKYDSMVOSLEDVTAQFESYKALTASETEDKLENS 480
 DB 421 LKGRKAELEKSSAAHTQATLLLOEKYDSMVOSLEDVTAQFESYKALTASETEDKLENS 480

QY 481 LOEKAAGKNAEDVQHILATESSNOEYVRMLLDLQTSALKETEIKEITVSFLQKITD 540
 DB 481 LOEKAAGKNAEDVQHILATESSNOEYVRMLLDLQTSALKETEIKEITVSFLQKITD 540
 QY 541 LONQLKQOEEDFRKOLEDEEGRKAKEKNTTAEITBEINKWRLLYBELYNKTPFQQLDA 600
 DB 541 LONQLKQOEEDFRKOLEDEEGRKAKEKNTTAEITBEINKWRLLYBELYNKTPFQQLDA 600
 QY 601 FEVEKQALLNEHGAQEQOLNKIRDSYAKLLGHQNLKQIKKHVVKLKDNSQKSEVSKLR 660
 DB 601 FEVEKQALLNEHGAQEQOLNKIRDSYAKLLGHQNLKQIKKHVVKLKDNSQKSEVSKLR 660
 QY 661 COLAKKQKQSETKLOELNKLKGIKHFDPKAPHHESKENFALKTPKLGKNTNRYAPMBC 720
 DB 661 COLAKKQKQSETKLOELNKLKGIKHFDPKAPHHESKENFALKTPKLGKNTNRYAPMBC 720
 QY 721 QESWK 725
 DB 721 QESWK 725
 RESULT 2
 ABG60842
 ID ABG60842 standard; Protein; 725 AA.
 AC ABG60842;
 XX
 DT 13-AUG-2002 (first entry)
 DE
 DE Human receptor for hyaluronan-mediated motility (RHAMM).
 KW Tissue disorder; response-to-injury process; cell proliferating;
 KW hyaluronic acid; HA; receptor for hyaluronan-mediated motility;
 KW RHAMM; inflammatory neurological disorder; Parkinson's disease;
 KW Alzheimer's disease; arthritis; multiple sclerosis; gastritis; nephritis;
 KW stenosis; restenosis; cancer; kidney fibrosis; inflammatory lung disease;
 KW emphysema; asthma; cystic fibrosis; obesity; obesity related disease;
 KW lupus; cardiovascular disease; atherosclerosis; wound; scar; diabetes;
 KW tissue transplantation; stroke; inflammatory response; fibrotic response;
 KW medical implant; Acquired immunodeficiency syndrome; AIDS; hepatitis;
 KW myocardial fibrosis; hepatic fibrosis; chronic cystitis; acute mastitis;
 KW septic shock; thyroiditis; retinopathy.
 XX Homo sapiens.
 OS
 XX WO200228415-A1.
 PN
 XX 11-APR-2002.
 PD
 XX 05-OCT-2000; 2000WO-IB01534.
 XX
 XX 05-OCT-2000; 2000WO-IB01534.
 XX
 XX (TRAN-) TRANSITION THERAPEUTICS & DIAGNOSTICS IN.
 XX
 XX Turley EA, Cruz TF;
 PI
 XX WPI; 2002-435298/46.
 DR
 XX
 XX Treating tissue disorder associated with response-to-injury process or
 XX proliferating cells in mammals, e.g. fibrosis, inflammation, by
 XX administering a compound that alters activity of transition molecules
 XX within a cell -
 XX
 XX Example 30; Fig 50; 215pp; English.
 PS
 XX
 XX The invention describes a method of treating a tissue disorder associated
 CC with response-to-injury process or proliferating cells in a patient,
 CC comprising administering a polypeptide (I) which binds hyaluronic acid
 CC (HA), an antibody which binds one of domains D1-D5 of Receptor for
 CC hyaluronan-mediated motility (RHAMM), a polypeptide fragment encoding

CC any of DI-D5 of RHAMM, or a vector which expresses antisense RHAMM,
CC antibodies or a polypeptide fragment. The method is useful for treating a
CC patient with an inflammatory neurological disorder such as Parkinson's
CC disease, Alzheimer's disease, arthritis including rheumatoid arthritis,
CC osteoarthritis, multiple sclerosis, inflammatory dermatosis (psoriasis),
CC inflammation bowel disease, stenosis or restenosis, cancer, kidney
CC fibrosis, inflammatory lung disease (e.g. emphysema, asthma, cystic
CC fibrosis), obesity or obesity related diseases, lupus, cardiovascular
CC disease (e.g. atherosclerosis), and wound especially surgical excision
CC adhesions, to prevent scar and also for treating or preventing diabetes
CC mellitus. The method is also useful for treating tissue transplantation
CC (e.g. skin grafts), stroke, inflammatory responses or fibrotic response
CC associated with medical implants such as hip implants, vascular wraps and
CC catheters), inflammatory diseases such as AIDS, myocardial and hepatic
CC fibrosis, chronic cystitis, acute mastitis, gastritis, nephritis,
CC hepatitis, septic shock, thyroiditis, and retinopathy. This sequence
CC represents a receptor for hyaluronan-mediated motility protein used in
CC the method of treating a tissue disorder described in the invention.
XX
SQ Sequence 725 AA;

Query Match 100.0%; Score 3598; DB 23; Length 725;
Best Local Similarity 100.0%; Pred. No. 5.3e-185;
Matches 725; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSFPKAPLKRENDSPGAPSGYDVKTLEVLKGPVFSQKQKESQKQNLNVKDT 60
Db 1 MSFPKAPLKRENDSPGAPSGYDVKTLEVLKGPVFSQKQKESQKQNLNVKDT 60
Qy 61 TLPASARKVKSSESQKNDKDLKILEKEIRVLLQERGAQDRRIQDLETELEKMEARL 120
Db 61 TLPASARKVKSSESQKNDKDLKILEKEIRVLLQERGAQDRRIQDLETELEKMEARL 120
Qy 121 NAALREKTSLSANNATLEKQIETRLNELLKSKFSENGNKNLRIISLEMLKLNRET 180
Db 121 NAALREKTSLSANNATLEKQIETRLNELLKSKFSENGNKNLRIISLEMLKLNRET 180
Qy 181 KMRGMMAKQEGEMKLVQTRSLSEESQKIAQLGKGLVSIKEKIDSEKSEKLVETEE 240
Db 181 KMRGMMAKQEGEMKLVQTRSLSEESQKIAQLGKGLVSIKEKIDSEKSEKLVETEE 240
Qy 241 ISCASDQVKEYKLDIALEENLKEKNDIISLKSLEENIVILSKQVEDLVNKQLEKE 300
Db 241 ISCASDQVKEYKLDIALEENLKEKNDIISLKSLEENIVILSKQVEDLVNKQLEKE 300
Qy 301 KEDHVNRRHNENLNAEMQNLKQFLEQOEHEKLOQKELQIDSLQOEKELSSSLHQK 360
Db 301 KEDHVNRRHNENLNAEMQNLKQFLEQOEHEKLOQKELQIDSLQOEKELSSSLHQK 360
Qy 361 LCSFQEMVKEKNLFEBELKQTLDELQKQEQAEERLVKQLEEEAKSRAEELKLLLEK 420
Db 361 LCSFQEMVKEKNLFEBELKQTLDELQKQEQAEERLVKQLEEEAKSRAEELKLLLEK 420
Qy 421 LKGEAELEKSSAAHTQATLLQEKYSMVQSLVEDVTAQFESYKALFTASEIEDLKLENS 480
Db 421 LKGEAELEKSSAAHTQATLLQEKYSMVQSLVEDVTAQFESYKALFTASEIEDLKLENS 480
Qy 481 LOEAAKAGNAEDVQHOILATESNQEYVRMLDLQTKSALKETEKEIVSVLQKITD 540
Db 481 LOEAAKAGNAEDVQHOILATESNQEYVRMLDLQTKSALKETEKEIVSVLQKITD 540
Qy 541 LQNLKQOEEDFRKQLEDEGRKAEKENTTAELTEENKRWLLYEELYNKTKPQLOLDA 600
Db 541 LQNLKQOEEDFRKQLEDEGRKAEKENTTAELTEENKRWLLYEELYNKTKPQLOLDA 600
Qy 601 FEVEKQALLNEHGAQEQNLKIRDSYAKLLGHQNLKQIKIHVVKLKDENSQKSEVSKLR 660
Db 601 FEVEKQALLNEHGAQEQNLKIRDSYAKLLGHQNLKQIKIHVVKLKDENSQKSEVSKLR 660
Qy 661 QOLAKKQSETKLOEELNKVLGIKHFDPDSKAFHHESKENFALKTPLEKGNVNCYRAPMEC 720
Db 661 QOLAKKQSETKLOEELNKVLGIKHFDPDSKAFHHESKENFALKTPLEKGNVNCYRAPMEC 720

Qy 721 QESWK 725
Db 721 QESWK 725
RESULT 3
ID AAU11436 standard; Protein; 725 AA.
XX AAU11436;
AC AAU11436;
XX 12-MAR-2002 (first entry)
XX Human hyaluronic acid binding protein RHAMM.
XX Human; hyaluronic acid binding protein; RHAMM; gene therapy;
KW receptor for HA mediated mobility; immunosuppressive; cytostatic.
KW conjugate; rheumatoid arthritis; scleroderma; liver fibrosis; cancer;
XX Homo sapiens.
XX Key Location/Qualifiers
FT Misc-difference 278 /note= "Encoded by GAA"
FT Misc-difference 299 /note= "Encoded by AAA"
FT Misc-difference 323 /note= "Encoded by AAA"
FT Misc-difference 331 /note= "Encoded by CAG"
XX WO200180899-A2.
XX 01-NOV-2001.
XX 20-APR-2001; 2001WO-CA00533.
XX 20-APR-2000; 2000US-198613P.
XX (CANG-) CANGENE CORP.
XX Woloski BMR, Williams AM, Sereda TJ, Wiebe DJ;
XX WPI: 2002-075094/10.
XX N-PSDB; AAS17496.
XX Protein conjugates that selectively target certain tissues and organs
XX useful for treating and preventing various diseases, comprises
XX glucose-aminoglycan-targeting domain conjugated to a therapeutic
XX protein.
XX Claim 6; Page 115; 121pp; English.
XX The invention relates to a conjugate comprising an hyaluronic acid (HA)
XX -binding protein e.g. RHAMM (receptor of HA mediated mobility) or peptide
XX contiguous with, or coupled to a polypeptide conjugated to a therapeutic
XX agent, and the polynucleotides encoding them. Also included is a method
XX for preparation of the HA-binding protein by inserting a first nucleotide
XX sequence encoding a HA-binding protein directly linked to a second
XX nucleotide sequence encoding a therapeutic protein into a suitable
XX vector, expressing the vector in an acceptable host, purifying conjugate
XX molecule from host or expression medium. The composition is useful for
XX altering in vivo the distribution of a therapeutic agent comprising
XX administering the composition to the animal where conjugate molecule will
XX distribute primarily in tissues and organs containing high levels of
XX endogenous HA and for treating mammal with a disorder where a diseased
XX tissue of the mammal contains high level of HA e.g. rheumatoid
XX arthritis, scleroderma, liver fibrosis and cancer. Lower therapeutic
XX dosages required also translates into lower immunogenicity of the
XX conjugated protein as compared to the native protein. As a result,
XX conjugates improve patient compliance and reduce direct and indirect
XX costs associated with the drug substance and its administration.
XX Conjugates allows for the use, where appropriate, of lower, safer,

CC dosages as compared to the conventional dosage requirements for the
CC unconjugated corresponding therapeutic agent. Conjugate molecules have an
CC increased half-life and potency, resulting in prolonged circulation of
CC the molecule, efficient distribution into the target tissues, and
CC increased bioavailability. The present sequence represents a RHAMM
CC protein.

XX	Sequence	725 AA;			
SO	Query Match	99.5%;	Score 3581;	DB 23;	Length 725;
	Best Local Similarity	99.4%;	Pred. No. 4.3e-184;		
	Matches 721;	Conservative	3;	Mismatches	1; Indels 0; Gaps 0;
Qy	1	MSFPKAPLKRENDPSGCAPSGAYDVKTLEVLKGPVSPQSKQRFKQKESQNINVDKDT	60		
Db	1	MSFPKAPLKRENDPSGCAPSGAYDVKTLEVLKGPVSPQSKQRFKQKESQNINVDKDT	60		
Qy	61	TLPASARKVKSSEKSKQNDKDLKILEKEIRVLLQERGAQDRRIQDLETELEKMEARL	120		
Db	61	TLPASARKVKSSEKSKQNDKDLKILEKEIRVLLQERGAQDRRIQDLETELEKMEARL	120		
Qy	121	NAALREKTSLSANNATLEKQIELTRTNELLSKSFSENGNKNLRILSLMLKLRNKRET	180		
Db	121	NAALREKTSLSANNATLEKQIELTRTNELLSKSFSENGNKNLRILSLMLKLRNKRET	180		
Qy	181	KMRGMAKQEGEMKQVQTSLSSESGKIAQLEGKLVYSIEKEKIDSEKLEVEYEE	240		
Db	181	KMRGMAKQEGEMKQVQTSLSSESGKIAQLEGKLVYSIEKEKIDSEKLEVEYEE	240		
Qy	241	ISCSADQVEKYKLDIALEENLEKNDILSKQSLEENIVILSKQVEDLVNKKOLEKE	300		
Db	241	ISCSADQVEKYKLDIALEENLEKNDILSKQSLEENIVILSKQVEDLVNKKOLEKE	300		
Qy	301	KEDVNRREHNENINAEWNLKQFIEQQEHEKLOQKELQIDSLQOKEKELSSSLHOK	360		
Db	301	KEDVNRREHNENINAEWNLKQFIEQQEHEKLOQKELQIDSLQOKEKELSSSLHOK	360		
Qy	361	LCSFOEMVMEKNLFEELKQTLDELQKQKEQAERLVKQLEEEAKSRAEELKLEEK	420		
Db	361	LCSFOEMVMEKNLFEELKQTLDELQKQKEQAERLVKQLEEEAKSRAEELKLEEK	420		
Qy	421	LKGKAELEKSSAAHTQTLILLOEKYDSMWQSLVEDVTAQFESYKALTASEIEDLKLENS	480		
Db	421	LKGKAELEKSSAAHTQTLILLOEKYDSMWQSLVEDVTAQFESYKALTASEIEDLKLENS	480		
Qy	481	LOEKAAKAGNAEDVHOILATESSNOEYVRMLDLQTKSALKETEIKETIVSFLOKITD	540		
Db	481	LOEKAAKAGNAEDVHOILATESSNOEYVRMLDLQTKSALKETEIKETIVSFLOKITD	540		
Qy	541	LQNQLKQOEEDFRKQLEDEGRKAKEKNTTAEITTEINKWRLLYEELYNKTKPFQQLDA	600		
Db	541	LQNQLKQOEEDFRKQLEDEGRKAKEKNTTAEITTEINKWRLLYEELYNKTKPFQQLDA	600		
Qy	601	FEVEKQALLNEHGAQEQNLKIRDSYAKLGHQNLKQIKHVKVLLKDNSOLKSEVSKLR	660		
Db	601	FEVEKQALLNEHGAQEQNLKIRDSYAKLGHQNLKQIKHVKVLLKDNSOLKSEVSKLR	660		
Qy	661	COLAKKKOSETKLOBELKNVLGIKHFDPSKAFHHESKENFALKTPLKEGNTNCRAPMEC	720		
Db	661	COLAKKKOSETKLOBELKNVLGIKHFDPSKAFHHESKENFALKTPLKEGNTNCRAPMEC	720		
Qy	721	QESWK 725			
Db	721	QESWK 725			

RESULT 4
AAU11438
ID AAU11438 standard; Protein: 713 AA.
XX
AC AAU11438;
XX
DT 12-MAR-2002 (first entry)

XX	Rat hyaluronic acid binding protein RHAMM.				
DE	Rat; hyaluronic acid binding protein; RHAMM; gene therapy;				
KW	receptor for HA mediated mobility; immunosuppressive; cytostatic.				
KW	conjugate; rheumatoid arthritis; scleroderma; liver fibrosis; cancer;				
XX	Rattus norvegicus.				
OS					
XX	WO200180899-A2.				
PN					
XX	01-NOV-2001.				
PD					
XX	20-APR-2001; 2001WO-CA00533.				
PF					
XX	20-APR-2000; 2000US-198613P.				
PR					
XX	(CANG-) CANGENE CORP.				
PA					
XX	Woloski BMR, Williams AM, Sereda TJ, Wiebe DJ;				
PI					
XX	WPI: 2002-075094/10.				
DR	N-PSDB; AAS17498.				
DR					
XX	Protein conjugates that selectively target certain tissues and organs				
PT	useful for treating and preventing various diseases, comprises				
PT	glucose-aminoglycan-targeting domain conjugated to a therapeutic				
PT	protein -				
XX	Claim 6; Page 116-117; 121pp; English.				
PS					
CC	The invention relates to a conjugate comprising an hyaluronic acid (HA)				
CC	-binding protein e.g. RHAMM (receptor of HA mediated mobility) or peptide				
CC	contiguous with, or coupled to a polypeptide conjugated to a therapeutic				
CC	agent, and the polynucleotides encoding them. Also included is a method				
CC	for preparation of the HA-binding protein by inserting a first nucleotide				
CC	sequence encoding a HA-binding protein directly linked to a second				
CC	nucleotide sequence encoding a therapeutic protein into a suitable				
CC	vector, expressing the vector in an acceptable host, purifying conjugate				
CC	molecule from host or expression medium. The composition is useful for				
CC	altering in vivo the distribution of a therapeutic agent comprising				
CC	administering the composition to the animal where conjugate molecule will				
CC	distribute primarily in tissues and organs containing high levels of				
CC	endogenous HA and for treating mammal with a disorder where a diseased				
CC	tissue of the mammal contains high level of HA e.g. rheumatoid				
CC	arthritis, scleroderma, liver fibrosis and cancer. Lower therapeutic				
CC	dosages required also translates into lower immunogenicity of the				
CC	conjugated protein as compared to the native protein. As a result,				
CC	costs associated with the drug substance and its administration.				
CC	Conjugates allows for the use, where appropriate, of lower, safer,				
CC	dosages as compared to the conventional dosage requirements for the				
CC	unconjugated corresponding therapeutic agent. Conjugate molecules have an				
CC	increased half-life and potency, resulting in prolonged circulation of				
CC	the molecule, efficient distribution into the target tissues, and				
CC	increased bioavailability. The present sequence represents a RHAMM				
CC	protein.				
XX					
SO	Sequence 713 AA;				

Query Match	74.2%;	Score 2670.5;	DB 23;	Length 713;
Best Local Similarity	75.3%;	Pred. No. 2.6e-135;		
Matches 537;	Conservative	71;	Mismatches 104;	Indels 1; Gaps 1;
Qy	1	MSFPKAPLKRENDPSGCAPSGAYDVKTLEVLKGPVSPQSKQRFKQKESQNINVDKDT	60	
Db	1	MSFPKAPLKRENDPSGCAPSGAYDVKTLEVLKGPVSPQSKQRFKQKESQNINVDKDT	60	
Qy	61	TLPASARKVKSSEKSKQNDKDLKILEKEIRVLLQERGAQDRRIQDLETELEKMEARL	120	
Db	61	TLPASARKVKSSEKSKQNDKDLKILEKEIRVLLQERGAQDRRIQDLETELEKMEARL	120	
Qy	121	NAALREKTSLSANNATLEKQIELTRTNELLSKSFSENGNKNLRILSLMLKLRNKRET	180	

Db 121 NAPVTEKPSLSANNASLEKRLTELTRANELKSKFSEDAHQNMRTLSLEMLKLNKRET 180
Qy 181 KMRGMAKQEGMEKMQVOTORSLEESQKIAQLEGKLVSIIEKIDSEKSETEKLEYIEE 240
Db 181 KMRGMAKQEGMEKMQVOTORSLEESQKIAQLEGKLVSIIEKIDSEKSETEKLEYIEE 240
Qy 241 ISCASDQVEKYKLDIAQLEENLKEKNDKILSKQSLSEENIVILSKQVEDLNVKQCOLLEKE 300
Db 241 ISCASDQVEKYKLDIAQLEENLKEKNDKILSKQSLSEENIVILSKQVEDLNVKQCOLLEKE 300
Qy 301 KEDHVNRRNHNENLNAMQNLKQKFILEQOQHEKHLQKQLEEDLNVKQCOLLEKE 360
Db 300 RDDLVSKDRRAESLSAEMQVLTKEKLLEROEYKLOQNELQSQSLQOQKELSAHLQOQ 359
Qy 361 LCSFQEEVMKKNLFEELKOTLDELKLOKEQOAEPLVKQLEEEAKSRAEELKLEEK 420
Db 360 LCSFQEEVMTSRNVFKEQLKLALDELDAVQKKEQSEKLVKQLEEEYKSTAEQURRDDL 419
Qy 421 LKGEAELEKSSAHTQATLLQKESYKALQVDSQVLESDVTAQFESYKALTASETEDLKLENS 480
Db 420 LREKEIELEKRTAAHAQVIAQEKYSDTAQTLDVTAQLESYKSSTLKEIEDLKLENI 479
Qy 481 LOEAAKAGKNAEDVQOILATESNQBYVRMLDLQTKSALKETEKEITVSLQKITD 540
Db 480 LOEKVMAEKREVDVQOQILTAESTNQBYAKVQDQNSSTLKAEITKEITSSYLEKITD 539
Qy 541 LQNLKQOQEDFRKOLEDEEGRKAENKNTTAELEENKRWLLYEELYNKTKPQLOQLDA 600
Db 540 LQNLKQOQEDFRKOLEDEEGRKAENKNTTAELEENKRWLLYEELYNKTKPQLOQLDA 600
Qy 601 FEVEKQALLNEHGAQEQNLKIRDSYAKLLGHQNLKQKIKHVVKLKDNSOLKSEVSKLR 660
Db 600 FEAEKQALLNEHGAQEQNLKIRDSYAKLLGHQNLKQKIKHVVKLKDNSOLKSEVSKLR 659
Qy 661 COLAKKQSETKQBELKNVIGIKHFDPSKAFHESKENFALKTPLKEGNTNC 713
Db 660 SOLAKRKQNELRQELGDKALGIRHFDPSKAFHESKENVLTPLKEGNTNC 712

RESULT 5
AAU11437
ID AAU11437 standard; Protein; 794 AA.
XX AC AAU11437;
XX DT 12-MAR-2002 (first-entry)
XX DE Mouse hyalurononic acid binding protein RHAMM.
XX KW Mouse; hyalurononic acid binding protein; RHAMM; gene therapy;
KW conjugate; rheumatoid arthritis; immunosuppressive; cytostatic.
XX OS Mus musculus.
XX FH Key Location/Qualifiers
FT Misc-difference 55 /note= "Encoded by AAC"
FT Misc-difference 71 /note= "Encoded by ACG"
FT Misc-difference 89 /note= "Encoded by CAA"
FT Misc-difference 91 /note= "Encoded by CAC"
FT Misc-difference 540 /note= "Encoded by ACT"
FT Misc-difference 668 /note= "Encoded by GAT"
XX FN WO200180899-A2.
XX PD 01-NOV-2001.

XX PF 20-APR-2001; 2001WO-CA005333.
XX PR 20-APR-2000; 2000US-198613P.
XX PA (CANG-) CANGENE CORP.
XX PI Wołoski BMR, Williams AM, Sereda TJ, Wiebe DJ;
XX WPI: 2002-075094/10.
XX N-PSDB: AAS17497.
XX Protein conjugates that selectively target certain tissues and organs
PT useful for treating and preventing various diseases, comprises
PT glucose-aminoglycan-targeting domain conjugated to a therapeutic
PT protein
XX Claim 6; Page 116; 121pp; English.
XX The invention relates to a conjugate comprising an hyalurononic acid (HA)
CC -binding protein e.g. RHAMM (receptor of HA mediated mobility) or peptide
CC contiguous with, or coupled to a polypeptide conjugated to a therapeutic
CC agent, and the polynucleotides encoding them. Also included is a method
CC for preparation of the HA-binding protein by inserting a first nucleotide
CC sequence encoding a HA-binding protein directly linked to a second
CC nucleotide sequence encoding a therapeutic protein into a suitable
CC vector, expressing the vector in an acceptable host, purifying conjugate
CC molecule from host or expression medium. The composition is useful for
CC altering in vivo the distribution of a therapeutic agent comprising
CC administering the composition to the animal where conjugate molecule will
CC distribute primarily in tissues and organs containing high levels of
CC endogenous HA and for treating mammal with a disorder where a diseased
CC tissue of the mammal contains high level of HA e.g. rheumatoid
CC arthritis, scleroderma, liver fibrosis and cancer. Lower therapeutic
CC dosages required also translates into lower immunogenicity of the
CC conjugated protein as compared to the native protein. As a result,
CC conjugates improve patient compliance and reduce direct and indirect
CC costs associated with the drug substance and its administration.
CC Conjugates allows for the use, where appropriate, of lower, safer,
CC dosages as compared to the conventional dosage requirements for the
CC unconjugated corresponding therapeutic agent. Conjugate molecules have an
CC increased half-life and potency, resulting in prolonged circulation of
CC the molecule, efficient distribution into the target tissues, and
CC increased bioavailability. The present sequence represents a RHAMM
XX protein.
XX SQ Sequence 794 AA;
Query Match 72.9%; Score 2623; DB 23; Length 794;
Best Local Similarity 68.0%; Pred. No. 1e-132;
Matches 542; Conservative 67; Mismatches 100; Indels 88; Gaps 3;
Qy 1 MSFPAKPLKRNNDPSGCAPSPGAYDVKYLEVLKVPVSFKSQSRFKQKESKONLVNDKDT 60
Db 1 MSFPAKPLKRNNDPSGCAPSPGAYDVKYLEVLKVPVSFKSQSRFKQKESKONLVNDKDT 60
Qy 61 TLPASARKVKSSEKESOKNDKDLKILEKEIRVLLQERGAODRRIDOLETELEKWEARL 120
Db 61 TLLASAKKAKKSVSKDKNDKDVLRLEKEIRALLQERGTODKRIQDMESLELEAKL 120
Qy 121 NAALREKTSLSANNATLEKQIETLRTTNELKSKFSENGNKNLRTLSLEMLKLNKRET 180
Db 121 NAAVREKTSLSANNATLEKQIETLRTTNELKSKFSENGNKNLRTLSLEMLKLNKRET 180
Qy 181 KMRGMAKQEGMEKMQVOTORSLEESQKIAQLEGKLVSIIEKIDSEKSETEKLEYIEE 240
Db 181 KMRGMAKQEGMEKMQVOTORSLEESQKIAQLEGKLVSIIEKIDSEKSETEKLEYIEE 240
Qy 241 ISCASDQVEKYKLDIAQLEENLKEKNDKILSKQSLSEENIVILSKQVEDLNVKQCOLLEKE 300
Db 241 ISCASDQVEKYKLDIAQLEENLKEKNDKILSKQSLSEENIVILSKQVEDLNVKQCOLLEKE 300
Qy 301 KEDHVNRRNHNENLNAMQNLKQKFILEQOQHEKHLQKQLEEDLNVKQCOLLEKE 360
Db 301 KEDHVNRRNHNENLNAMQNLKQKFILEQOQHEKHLQKQLEEDLNVKQCOLLEKE 360

XX
PN EP721012-A2.
PD 10-JUL-1996.
PF 16-OCT-1995; 95EP-0307310.
XX
PR 14-OCT-1994; 94GB-0020740.
XX
PA (MANI-) MANITOBA CANCER TREATMENT & RES FOUND.
PA (UYMA-) UNIV MANITOBA.
PI Entwistle J, Turley EA, Zhang S;
XX
DR WPI; 1996-310997/32.
DR N-PSDB; AAT34525.
XX
PT Receptor for hyaluronic acid-mediated motility protein, and DNA
PT encoding it - useful to treat or prevent diseases associated with
PT the receptor, e.g. breast cancer
XX
PS Claim 8; Page 50-52; 117pp; English.
XX
CC RHAMM 1-2a (AAR99675) is an alternatively spliced variant of
CC RHAMM 1 (AAR99673) (receptor for hyaluronic acid mediated
CC motility), a protein involved in cell locomotion or motility and
CC cell proliferation and transformation. It differs from RHAMM 1
CC by an insertion of 25 amino acids (see also AAR99674) between
CC amino acids 54 and 55 of RHAMM 1, resulting from an alternatively
CC spliced exon 2A (AAT34502). RHAMM 1-2a is the isoform that is
CC overexpressed in tumours. Determination of the level of RHAMM
CC 1-2a in a sample can be used to assess the prognosis of a tumour
CC (esp. breast cancer) patient. The RHAMM 1-2a protein can also
CC be used to suppress or control a tumour by modulating the
CC interaction of cell-associated RHAMM with its ligand.
XX
SQ Sequence 631 AA;

Query Match 54.3%; Score 1952; DB 17; Length 631;
Best Local Similarity 64.7%; Pred. No. 7e-97;
Matches 410; Conservative 51; Mismatches 85; Indels 88; Gaps 3;

QY 164 LRILSLEMLKLNKRETKMRGMMAQOEGEMKLVQVRSLESOQKIAQLEGKLVISKE 223
DB 1 MRALSLEMLKLNKRETKMRSMVMVQEGMELQATQKDLTESKIVQLEGKLVISKE 60
QY 224 KIDKSETEKLEYTEETSCASDOVEKYKLDIAQLEENLKEKNDRIILSKQSLEENIVIL 283
DB 61 KIDKCEKLEYTOEISCASDQVEKQVDAQLEEDLKEKDRILSLKQSLEENIT-F 119
QY 284 SKQVEDLVKCOLLEKEDHVNRRNHNENLNAEMQNLKOFLEQOEHEKLOQKELOI 343
DB 120 SKQIEDLVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALERQYKLOQKELOQS 179
QY 344 DLSLOQKEKLSLHQKLCFQOEMVMEKKNLFEELKOTLDLDELKLOQKEQAEKLVKOL 403
DB 180 QSLLOQKEKLSARLOQQLCSFOEEMTSEKNVFEKELKALAEALDAVQOKEQSERLVKOL 239
QY 404 EEEAKSRAEELKLEKLGKAEAEKSAHAQTATL----- 441
DB 240 EEEKSTAEQLTRLDNLLREKEVELEKHIAHAQAAILIAQEKYNDTAQSLRDVTAQLESV 299
QY 442 ----- 441
DB 300 QEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQSL 359
QY 442 --LQEKYDSMVQSLDVTVAQFESYKALPASETDLKLNSLSQEKAAKAGNAEDVQHOI 499
DB 360 ESVQEKYNDTAQSLRDVTAQLESYKSSSTLKEIEDLKLNLTLQEKVAMAEKSVEDVQOOI 419
QY 500 LATESNOEYVRLMDLDTKSKALTEKEITVSVFLQITDLQNLKQOEDFRKQLEDE 559
DB 420 LTAESTNOEYARMVODLQNRSTLKEEIKETSSFLEXITDLKNQLROQDEDFRKLQLEEK 479

QY 560 EGRKAKEKNTAELTEINKWRLLYEELYNKTKPFQLQDAFEVEKQALLNEHGAQEQOL 619
DB 480 GKRTAEKENVTMETLMEINKWRLLYDELYEKTTPQOQDAFEAEKQALLNEHGATQEQOL 539
QY 620 NKIRDSYAKLLGHQNLKOKIKHVVKLKDENSQKSEVSKLRCQLAKKQSETKLOEELNK 679
DB 540 NKIRDSYAQLLGHQNLKOKIKHVVKLKDENSQKSEVSKLRSQSLVKKRQNLRLQGEIDK 599
QY 680 VLGKHKHEDPSKAFHESKENFALKTPLKEGNTNC 713
DB 600 ALGIRHFDPSKAFCHASKENF---TPLKEGNTNC 630

RESULT 9
AAR99673
ID AAR99673 standard; Protein; 606 AA.
XX
AC AAR99673;
XX
DT 10-OCT-1996 (first entry)
XX
DE Receptor for hyaluronic acid mediated motility RHAMM 1.
XX
KW RHAMM 1; receptor for hyaluronic acid mediated motility;
KW hyaluronan receptor; cell locomotion; cell proliferation;
KW breast cancer; therapy.
XX
OS Mus sp.
FH Key Location/Qualifiers
FT Modified-site 91..93
FT Modified-site 258..260 /label= N-glycosylation_site
FT Modified-site 279..281 /label= N-glycosylation_site
FT Modified-site 300..302 /label= N-glycosylation_site
FT Modified-site 321..323 /label= N-glycosylation_site
FT Modified-site 342..344 /label= N-glycosylation_site
FT Modified-site 373..375 /label= N-glycosylation_site
FT Modified-site 413..415 /label= N-glycosylation_site
FT Domain 532..542 /label= Hyaluronan_binding_domain-I
FT Domain 553..562 /label= Hyaluronan_binding_domain-I
FT Modified-site 594..596 /label= N-glycosylation_site
XX
PN EP721012-A2.
XX
PD 10-JUL-1996.
XX
PF 16-OCT-1995; 95EP-0307310.
XX
PR 14-OCT-1994; 94GB-0020740.
XX
PA (MANI-) MANITOBA CANCER TREATMENT & RES FOUND.
PA (UYMA-) UNIV MANITOBA.
PI Entwistle J, Turley EA, Zhang S;
XX
DR WPI; 1996-310997/32.
DR N-PSDB; AAT34499;
DR N-PSDB; AAT34500;
DR N-PSDB; AAT23423;
DR N-PSDB; AAT34524.
XX

Qy	439	TLL-----	44			
Db	121	ILIAQEKYIDTAOSLRVYVTAQLESVQOEKYNDTAQSLRDVTAQLESQOEKYNDTAQSLRDV	180			
Qy	442	-----LQEKYDSMWQSLDVTDAQESYKALFTASIEDL	474			
Db	181	TAQLESQOEKYNDTAQSLRDVTAQLESVQOEKYNDTAQSLRDVTAQLESYKSSTLKEIEDL	240			
Qy	475	KLENSLQEKAAKAGNAEDVHQHILATESNQEYVRMLDLQTKSALKETEIKETIVSF	534			
Db	241	KLENTLQEKVAMAEKSVEDVQQILTAESTNQEYARMVQDLQNRSTLKEEIKEITSSF	300			
Qy	535	LOKTTDLQNLQOEEDFRKQLEDEGRKAENKNTAELTEETINKWRLLYEELYNTKPF	594			
Db	301	LEKITDLKNLRQOODEFRKQLEBKGRRTAEKENVMVTELTMEINKWALLY-ELYEKTKEP	359			
Qy	595	QLOLDAFEVEKQALLNEHGAQOLNKIRDSYAKLLGHQNLKOKIKHVVKLKDENSOLKS	654			
Db	360	QQQUDAFEAKQALLNEHGAQOLNKIRDSYAKLLGHQNLKOKIKHVVKLKDENSOLKS	419			
Qy	655	EVSRLQOLAKKKQSEPKLQELNKKVLGIRKHPDSPKAFHSHESKENFALKTPLKEGNTNC	713			
Db	420	EVSRLQOLVKKRQNELRLQOGLDKALGIRHEDPSRAFCASKENE--TPLKEGPNPC	475			
RESULT 12						
AAW02258						
ID	AAW02258 standard; Protein; 1411 AA.					
XX						
AC	AAW02258;					
XX						
DT	09-MAR-1997 (first entry)					
XX						
DE	Nucleolar/endosomal auto-antigen p162.					
XX						
KW	Auto-antibody; p162; rheumatic disease; antigen; diagnosis;					
KW	gene therapy.					
XX						
OS	Homo sapiens.					
XX						
PN	DE19515314-Cl.					
XX						
PD	12-SEP-1996.					
XX						
PF	27-APR-1995; 95DE-1015514.					
XX						
PR	27-APR-1995; 95DE-1015514.					
XX						
PA	(PRIV-) PRIVATES INST IMMUNOLOGIE & MOLEKULARGEN.					
XX						
PI	Renz M, Seelig HP;					
XX						
DR	WPI; 1996-403153/41.					
DR	N-PSDB; AAT58751.					
XX						
PT	DNA encoding nucleolar-endosomal auto-antigen - useful for exact					
PT	diagnosis of rheumatic disease, in gene therapy and for removal of					
PT	specific auto-antibodies					
XX						
PS	Claim 1; Fig 2; 15pp; German.					
XX						
CC	Transformed cells can be cultured to produce the antigen p162, for use					
CC	in exact (differential) diagnosis of rheumatic disease, i.e. they					
CC	can detect, in immunoassays, Western blots, etc., rheumatism-					
CC	specific auto-antibodies. The antigen can be used therapeutically,					
CC	in the removal of auto-antibodies from the circulation, or when					
CC	coupled to a cytotoxin, the elimination of auto-antibody-					
CC	producing lymphocytes.					
XX						
SQ	Sequence 1411 AA;					
Query Match 11.9%; Score 429.5; DB 17; Length 1411;						
Best Local Similarity 25.2%; Pred. No. 6.3e-15;						

QY 253 LDTAQLEENLKEKNDEILSLKOSLEE-----NIVILSKQVEDLVNVCQLEKEKEDHV 305
Db 1379 LALQSLADTKKKVDDDLGTTIESLEAKKLLKDAEALQSLEEKALAYDKLETK---- 1434
QY 306 NNREHNENLNAEMQNLKQKFILEQOEHEKLOQKELOLDSLLQOEKELSSSLHQR----- 360
Db 1435 NRLQOEELDTVTDLDH-----QRVASNLKQKQKFDOLLAEBKSIARYAEERDRAE 1487
QY 361 -----LCSFOEMVKEKNLFEELKQTLDELKQKEQAEERLVKQLEEBK 408
Db 1488 AEAREKETKALSARALEAEAKFEFERQNKQLRADHEDLMSKDDVGVKNVHELEKSKR 1547
QY 409 SRAEELKLEELKQKEAELEKSSAAHTQATLLQEKYDSMVQSLDVTQAQFESYKALTA 468
Db 1548 ALEQQVEEMRTQLEELDELEQATDAKLRLVNMQAMKAQFEDLQTRDEQNEKKRLLI 1607
QY 469 SEIEDL--KLENSSLOEKAAKAGK-----NAEDVOHQILATESSNQEVYVRLMLDLQTKSA 521
Db 1608 KQVRELEAELEDERKORALAVASKKKMEIDLKDLAQIEAANKARDEVIKQLRKLQAQMK 1667
QY 522 LKETEKETTVS---FLQ-----KITDLQNLKQOE-----EDFRKQLED 558
Db 1668 DYQRELEEARSRDIFAQSEKSEKLLKSLAEIILQOEELASSRRARRHAEQERDELAD 1727
QY 559 E-----EGRKA---EK---ENTTAELTEENKRWLLYEELNKTTPFQQLDAPFEVQQA 607
Db 1728 EITNSASGKSALLDEKRLREARIAQLEEELEEQSNMELLNDRFRKTTLVQVDTLNAE--- 1784
QY 608 LLNEHGAQOEQLNKTRDSYAKLLGHON--LKQIKHV-----VKLKDENSQKSEVSKLR 660
Db 1785 LAAERSAAQKS-----DNARQOOLERONKELKAKQLEGAVKSKFKATISALEAKIGOLE 1839
QY 661 COL---AKKQSETKLOEELNKVL-----GIKFDPSPKAFPHSKENF-ALKTPLK 707
Db 1840 EQLEQAEKERAANKLVRTKELKEIFQVEDERRHADQYKEQMEKANARKQLKROLE 1899
QY 708 EGNWNCYRA 716
Db 1900 EAEEEAATRA 1908
RESULT 14
AAW41000
ID AAW41000 standard; Protein; 1988 AA.
AC AAW41000;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 5931.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.

29-NOV-2000; 2000US-0727344.
(HYSE-) HYSEQ INC.
Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
Zhao QA, Zhou P, Goodrich R, Drmanac RT;
WPI: 2001-442253/47.
N-PSDB; AAI60156.
Novel nucleic acids and polypeptides, useful for treating disorders
such as central nervous system injuries -
Example 2; SEQ ID NO 5931; 10078pp; English.
The invention relates to human nucleic acids (AAI57798-AAI61369) and
the encoded polypeptides (AAM38642-AAM42213) with nootropic,
immunosuppressant and cytostatic activity. The polynucleotides are useful
in gene therapy. A composition containing a polypeptide or polynucleotide
of the invention may be used to treat diseases of the peripheral nervous
system, such as peripheral nervous injuries, peripheral neuropathy and
localised neuropathies and central nervous system diseases, such as
Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX Sequence 1988 AA;
SQ
Query Match 11.5%; Score 414; DB 22; Length 1988;
Best Local Similarity 23.4%; Pred. No. 6.2e-14;
Matches 199; Conservative 150; Mismatches 310; Indels 190; Gaps 29;
QY 32 LKGPVSFOKSORFKOOKESKONLVNDKDTTL-PASARKV-----KSS 72
Db 1086 LQAQIDELKQLAKKEELQALARGDDETLHKNNALKVVRLEQQAIAELQEDFESEKAS 1145
QY 73 ESKKESQND--KDLKILEKEIRVLLQERGAQ-----DRR 105
Db 1146 RNKAEKQKRDLSSEELAKTELEDTLDTTAAQQLRTKREQEVAELKALEETKNHEAQ 1205
QY 106 IQDLE----TELEKMEARLNAALREKTSLSANNATLEKQELIETRTNELLKSKFSENGNQ 161
Db 1206 IQDMRQRHATALEELSEOLEQAKRFKANLEKNKQGLTDNNKELACEVKVLQOVKAESHK 1265
QY 162 K-----NLRI-LSLELMKLRNK-----RETKMRGMMAQGE--GM 192
Db 1266 RKKLDAQVQELHAKVSGDRULVELAEKASKLQNELONVSTLLEAEKKGKFKAKDAASL 1325
QY 193 EMKLQVTRQSLSEESQGIATQLEGKLVSTEKIDSEKTEKLEIYIEISCAQDQVEKYK 252
Db 1326 ESQLODTQELLQEEETROKLNLSRIRQLEEK-----NSLQEQEEEEARKNLEKQV 1378
QY 253 LDTAQLEENLKEKNDEILSLKOSLEE-----NIVILSKQVEDLVNVCQLEKEKEDHV 305
Db 1379 LALQSLADTKKKVDDDLGTTIESLEAKKLLKDAEALQSLEEKALAYDKLETK---- 1434
QY 306 NNREHNENLNAEMQNLKQKFILEQOEHEKLOQKELOLDSLLQOEKELSSSLHOK----- 360
Db 1435 NRLQOEELDTVTDLDH-----QRVASNLKQKQKFDOLLAEBKSIARYAEERDRAE 1487
QY 361 -----LCSFOEMVKEKNLFEELKQTLDELKQKEQAEERLVKQLEEBK 408
Db 1488 AEAREKETKALSARALEAEAKFEFERQNKQLRADHEDLMSKDDVGVKNVHELEKSKR 1547
QY 409 SRAEELKLEELKQKEAELEKSSAAHTQATLLQEKYDSMVQSLDVTQAQFESYKALTA 468

Db 1548 ALEQQVEEMRTQLEBELEDELOQADAKLRLEVNMQAMKAQFERDLQTRDEQNEEKRLLI 1607
 Qy 469 SEIEDL--KLENSISLQKAAKAGK----NAEDVOHQILATESNQEVYRMLLDLQTKSA 521
 Db 1608 KQVRELEAELEDERKQALAVASKKWEIDLKLEAQIEAANKARDEVIKRLKLAQMK 1667
 Qy 522 LKETEIKETIVS----FLQ-----KITDLQNLKQOE-----EDFRKQLED 558
 Db 1668 DYQRELEEARSRDEIFAQSEKSEKSLKLEAEILQLQOEELASSERARRHAEQERDELAD 1727
 Qy 559 E-----EGRKA---EK---ENTTAELTEINKWRLLYEELYNKTPQLQLDAREVEKQA 607
 Db 1728 EITNSAKSALLDEKRLREARIAQLEEELEEQSNMELLNDRKTTQLQVDTLNAE--- 1784
 Qy 608 LLNEHGAQEOLNIRDSYAKLLGHON--LKOKIKHV-----VKLDKENSOLKSEVSKLR 660
 Db 1785 LAERSAAQKS-----DNARQOLERQNKELKAKLOELEGAVKSKFKATISALEAKIQLE 1839
 Qy 661 CQL---AKKQSEYKLOEELNKVL-----GIKHFDPSKAPHHESKENF-ALKTPLK 707
 Db 1840 EQLEQEAERAAANKLVRRTEKKLKEIFMQVEDERRHADQYEQMEKANARMKQLKQLE 1899
 Qy 708 EGNTCYRA 716
 Db 1900 EABEATRA 1908

RESULT 15
 AAU79590
 ID AAU79590 standard; Protein; 1388 AA.
 XX AC AAU79590;
 XX DT 24-SEP-2002 (first entry)
 XX DE Human kinesin motor protein, Hskif15.
 XX KW Human; enzyme; cytosolic; neuroprotective; kinesin motor protein; Hskif15; microtubule stimulated ATPase activity; Hskif15M1A367;
 KW Hskif15M1A367; kinesin; mitotic spindle; microtubule motor; antigen; therapeutic; diagnosis; cancer; neurological disorder; vesicular transport; atherosclerosis; tumour; abnormal wound healing; inflammatory disorder; immune disorder; rheumatoid arthritis; ocular angiogenic disease; glaucoma; cardiovascular disease; hypertension; diastolic dysfunction; fungal disease; aspergilliosis.
 XX OS Homo sapiens.
 XX PN US6391613-B1.
 XX PD 21-MAY-2002.
 XX PF 27-NOV-2000; 2000US-0723219.
 XX PR 04-JUN-1999; 99US-137423P.
 XX PR 17-MAY-2000; 2000US-0572191.
 XX PA (CYTO-) CYTOKINETICS INC.
 XX PI Beraud C, Sakowicz R, Wood KW;
 XX DR WPI; 2002-498776/53.
 XX DR N-PSDB; ABK86978.
 XX PT New nucleic acid encoding a human kinesin motor protein designated Hskif15 which has microtubule stimulated ATPase activity, for diagnosing and treating cancer, neurological disorders and disorders of vesicular transport -
 XX PS Claim 1; Fig 2; 21pp; English.
 XX SS The invention discloses an isolated nucleic acid encoding a human kinesin motor protein, Hskif15, which has microtubule stimulated ATPase activity.

CC and two truncated versions, Hskif15M1A367 and Hskif15M1A401, retaining the amino terminus. Hskif15, a member of the kinesin superfamily, has been found to be essential for mitotic spindle formation. The predicted structure comprises an amino-terminal kinesin-like microtubule "motor" domain. The polynucleotides and polypeptides may be used therapeutically to diagnose and prevent or treat cancer, neurological disorders and disorders of vesicular transport. Examples of the conditions that can be treated include atherosclerosis, tumours, abnormal wound healing, inflammatory and immune disorders (such as rheumatoid arthritis), ocular angiogenic disease (such as glaucoma), cardiovascular disease (such as hypertension), diastolic dysfunction and fungal disease (such as aspergilliosis). The polynucleotides and polypeptides may also be used to screen for modulators of Hskif15 and raise antibodies. The sequence presented is the human kinesin motor protein, Hskif15.

XX SQ Sequence 1388 AA;
 Query Match 11.5%; Score 413; DB 23; Length 1388;
 Best Local Similarity 21.0%; Pred. No. 4.7e-14;
 Matches 171; Conservative 157; Mismatches 282; Indels 206; Gaps 25;

Qy 37 SFQSKSREFKQ-----QKESQNLDV-----K 58
 Db 598 SKQYEFEFKELTRKQLESELSQKANKLNLENLEATKACKRQEVSQLNKHAEITLK 657
 Qy 59 DTTLPASARKVYKSSSKES-----QKNDKDKILEK 90
 Db 658 IITPTKAYQLHSRPVKLSPMGSGSLYTONSILNDILNEPVPPEMNEQAEALISE 717
 Qy 91 EIRVLQERGAQDRIQD-----LETELEKME-----ARLNAAERKTSLS 131
 Db 718 ELRTVQEQMSALQAKLDEEHKLNKLQOHVDKLEHHSTQMQLFSSERIDWTQKQEBLLS 777
 Qy 132 ANNATLEKQLIELTFTNELLKSPSE-----NGQKNLRILSLELMKLRNRETMRGM 186
 Db 778 QLN-VLEKQLQETQTKNDFLKSEVHDLRVLHSAADKLSVSKVSEPKTQOERFNKL- 835
 Qy 187 AKQEGMEKMLQVQORSLE-----ESQ-----KIAQEGKLVSIIEKEID 226
 Db 836 -SERHMHVQLQDLNLRLENKLESKACLDQSDYNLQIMKEFIDQLSRNLFQFKENET 894
 Qy 227 EKSTETKLELYIEEISCASQDVYKLDIAQLEENLKEKNDEILSKOSLEENIVILSKQ 286
 Db 895 LKSDLNLMLELLE-----AEKERNKLSL-QFEEDKNSKEILLKVLAEVRO-----EKQ 943
 Qy 287 VEDLNVCQ-----LLEKEKE-DHVNRRNHNENLNAMQNLKQKFILEQOE 333
 Db 944 KE--TAKCEQOMAKVQKLEESLATEKVISLSKSRSDSKKVVADLMNQIQEL-----R 995
 Qy 334 EKLOQKELQIDSLQOEKELSSSLHOKLCSFQOEMVK----- 370
 Db 996 SSVCEKETITDLKQELADINKYNSALVDREESRVLLKQOEVDILDLKETLRLILSED 1055
 Qy 371 -EKNLFEELKOTLDLQKLOQKQAEQRYKQLEEEAKSRAEELKLEELKKGAELE 429
 Db 1056 IERDMCEDLAHATEQLNMLTEASKHSGLLQSQAEELTKKEALIQELQHLNOKKEVE 1115
 Qy 430 KSSAAHTQATILLQEKYDSMVQSLED--VTAQFESYKA---LTASEYEDLKLENSLQE 483
 Db 1116 QKNEYNFKMRQLEHVMDSAADPQSPKTPPHFQTHAKLLETQEQEEDGRASKTSLEH 1175
 Qy 484 KAAKAGNAEDVQHQILATESNQEVYRMLLDLQ---TKSALKETEIKEITVSFQKITD 540
 Db 1176 LVTKLNEDREVKNAILRMKEQLREMNLRLESQOLIEKNWLLQQLQDIDK---RQKENS 1232
 Qy 541 LQN-----QLKQOEEDFRKQ-----LEDEEGRKAERKENTFAELTEIN 578
 Db 1233 QDNHPDNQQLKNEQESIKERLAKSKIIVEEMKMKADLEEVSALYNKMECLRTWDEVE 1292
 Qy 579 KWRLLYEELYNKTPFQQLDAFEVEKQALLNEHGAQEOQLNKIRDSYAKLLGHONLKQK 638
 Db 1293 RTQTLSEKAFQEKQLRSLKEEMVEERERTSQEEMMLRKQVECLAEEGKLVGHQNLHOK 1352

Search completed: December 20, 2002, 15:11:17
Job time : 45.4263 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:06:53 ; Search time 9.30678 Seconds
(without alignments)
2812.095 Million cell updates/sec

Title: US-09-685-010-48

Perfect score: 3104

Sequence: 1 MRALSLEMLKLNKRETKMR.....FCHASKENFTPLKEGNPNC 631

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	DB ID	Description
1	3104	100.0	794	1	Q00547 mus musculus
2	2091.5	67.4	498	1	HMNR_MOUSE
3	1949	62.8	724	1	HMNR_RAT
4	431.5	13.9	1790	1	USO1_YEAST
5	421.5	13.6	1939	1	MYH4_HUMAN
6	421	13.6	1938	1	MYH4_RABIT
7	417.5	13.5	1941	1	MYH2_HUMAN
8	415	13.4	1939	1	MYH1_HUMAN
9	415	13.4	1940	1	MYH3_RAT
10	413	13.3	1938	1	MYSS_CHICK
11	408	13.1	1130	1	YL17_CAEEL
12	406	13.1	1935	1	MYSS_CYPCA
13	404	13.0	1940	1	MYH3_CHICK
14	400.5	12.9	1084	1	MYSS_RABIT
15	400	12.9	1940	1	MYH3_HUMAN
16	396	12.8	1938	1	MYHD_HUMAN
17	393	12.7	1939	1	MYH6_MESAU
18	389.5	12.5	1935	1	MYH7_HUMAN
19	386.5	12.5	1939	1	MYH6_HUMAN
20	387	12.5	1938	1	MYH6_MOUSE
21	387	12.5	1938	1	MYH6_RAT
22	384	12.4	1960	1	MYH9_HUMAN
23	383.5	12.4	1976	1	MYHA_RAT
24	383	12.3	1427	1	REST_HUMAN
25	382.5	12.3	1102	1	MYSC_CHICK
26	379.5	12.2	1937	1	MYH8_HUMAN
27	379.5	12.2	1938	1	MYSAEQIR
28	378.5	12.2	1934	1	MYH7_MESAU
29	377	12.1	1935	1	MYH7_RAT
30	376.5	12.1	1978	1	MYA0_AQUAE
31	375.5	12.1	1976	1	MYHA_HUMAN
32	375	12.1	2230	1	GOG4_HUMAN
33	373.5	12.0	1961	1	MYH9_RAT

RESULT 1

ID	HMNR_MOUSE	STANDARD;	PRT;	794 AA.
AC	Q00547;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	Hyaluronan mediated motility receptor (Intracellular hyaluronic acid binding protein) (Receptor for hyaluronan-mediated motility).			
GN	HMNR OR IHABP OR RHAMM.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
[1]				
RN	SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.			
RC	TISSUE=Lung;			
RX	MEDLINE=98264863; PubMed=9601097;			
RA	Hofmann M., Fieber C., Assmann V., Goettlicher M., Sleeman J.,			
RA	Plug R., Howells N., von Stein O., Ponta H., Herrlich P.;			
RT	"Identification of IHABP, a 95 kDa intracellular hyaluronate binding protein."			
RT	J. Cell Sci. 111:1673-1684(1998).			
[2]				
RN	SEQUENCE FROM N.A.			
RP	Zhao Y., Zhang S., Turley E.;			
RA	Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.			
[3]				
RN	SEQUENCE OF 1-183 FROM N.A., AND TISSUE SPECIFICITY.			
RC	STRAIN=129/SV;			
RC	MEDLINE=99107769; PubMed=9889313;			
RA	Fieber C., Plug R., Sleeman J., Dall P., Ponta H., Hofmann M.;			
RT	"Characterization of the murine gene encoding the intracellular hyaluronan receptor IHABP."			
-RL	Gene 226:41-50(1999).			
[4]				
RN	SEQUENCE OF 164-794 FROM N.A., AND ALTERNATIVE SPLICING.			
RP	STRAIN=BALB/c; TISSUE=Fibroblast;			
RC	MEDLINE=96011639; PubMed=7590272;			
RA	Entwistle J., Zhang S., Yang B., Wong C., Li Q., Hall C.L., A J.,			
RA	Mowat M., Greenberg A.H., Turley E.A.;			
RT	"Characterization of the murine gene encoding the hyaluronan receptor RHAMM."			
RL	Gene 163:233-238(1995).			
[5]				
RP	SEQUENCE OF 318-794 FROM N.A., AND CHARACTERIZATION.			
RC	STRAIN=BALB/c;			
RA	MEDLINE=92299690; PubMed=1376732;			
RA	Hardwick C., Hoare K., Owens R., Hohn H.P., Hook M., Moore D.,			
RA	Cripps V., Austen L., Nance D.M., Turley E.A.;			
RT	"Molecular cloning of a novel hyaluronan receptor that mediates tumor cell motility."			
-RL	J. Cell Biol. 117:1343-1350(1992).			
[6]				
RN	CHARACTERIZATION.			
RP	MEDLINE=94308286; PubMed=7518470;			
RX				

Q27991 bos taurus
P10587 gallus gall
P79293 sus scrofa
P35748 oryctolagus
O42184 gallus gall
Q99323 drosophila
P32380 saccharomyc
P54697 dictyosteli
Q02455 saccharomyc
P35749 homo sapien
P02566 caenorhabdi
Q15431 homo sapien

ALIGNMENTS

RA Hall C.L., Wang C., Lange L.A., Turley E.A.;
 RT "Hyaluronan and the hyaluronan receptor RHAMM promote focal adhesion
 RL turnover and transient tyrosine kinase activity."; J. Cell Biol. 126:575-588(1994).
 RN [7]
 RP ERK REGULATION, AND SUBCELLULAR LOCATION.
 RX MEDLINE=98225222; PubMed=9556628;
 RA Zhang S., Chang M.C., Zylka D., Turley S., Harrison R., Turley E.A.;
 RT "The hyaluronan receptor RHAMM regulates extracellular-regulated
 RL kinase."; J. Biol. Chem. 273:11342-11348(1998).
 RN [8]
 RP REVIEW.
 RA MEDLINE=99059494; PubMed=9845361;
 RX Hofmann M., Assmann V., Fieber C., Sleeman J.P., Moll J., Ponta H.,
 RA Hart I.R., Herrlich P.;
 RT "Problems with RHAMM: a new link between surface adhesion and
 RT oncogenesis?"; Cell 95:591-592(1998).
 RL Cell 95:591-592(1998).
 CC -1- FUNCTION: INVOLVED IN CELL MOTILITY. WHEN HYALURONAN BINDS TO
 CC HMHR, THE PHOSPHORYLATION OF A NUMBER OF PROTEINS, INCLUDING THE
 CC FOCAL ADHESION KINASE OCCURS. MAY ALSO BE INVOLVED IN CELLULAR
 CC TRANSFORMATION AND METASTASIS FORMATION, AND IN REGULATING
 CC EXTRACELLULAR-REGULATED KINASE (ERK) ACTIVITY.
 CC -1- SUBUNIT: SUBUNIT OF THE HARC COMPLEX.
 CC -1- SUBCELLULAR LOCATION: CELL SURFACE AND CYTOPLASMIC.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; RHAMM1V4 (SHOWN HERE) AND
 CC RHAMM1; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: UBIQUITOUSLY EXPRESSED.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF031932; AAC12655.1; -
 DR EMBL; AF079222; AAD08670.1; -
 DR EMBL; AJ005919; CAA06768.1; -
 DR EMBL; AJ005920; CAA06768.1; JOINED.
 DR EMBL; AJ005921; CAA06768.1; JOINED.
 DR EMBL; AJ005922; CAA06768.1; JOINED.
 DR EMBL; AJ005923; CAA06768.1; JOINED.
 DR EMBL; AJ005924; CAA06768.1; JOINED.
 DR EMBL; X64550; CAA45849.1; -
 DR EMBL; X64550; CAA45848.1; -
 DR MGD; MGI:104667; Hmhr.
 KW Hyaluronic acid; Alternative splicing; Repeat; Glycoprotein.
 FT DOMAIN 442 546 5 X 21 AA NEAR PERFECT TANDEM REPEATS.
 FT REPEAT 442 462 1.
 FT REPEAT 462 483 2.
 FT REPEAT 483 504 3.
 FT REPEAT 504 525 4.
 FT REPEAT 525 546 5.
 FT REPEAT 546 561 6.
 FT DOMAIN 719 729
 FT DOMAIN 741 750
 FT CARBOHYD 53
 FT CARBOHYD 134
 FT CARBOHYD 279
 FT CARBOHYD 446
 FT CARBOHYD 467
 FT CARBOHYD 488
 FT CARBOHYD 509
 FT CARBOHYD 530
 FT CARBOHYD 561
 FT CARBOHYD 601
 FT VARSPLIC 218
 FT CONFLICT 19
 FT CONFLICT 55
 FT CONFLICT 71
 FT CONFLICT 89
 EKE -> OKH (IN REF. 2).

FT CONFLICT 94 94 A -> V (IN REF. 3).
 FT CONFLICT 540 540 S -> T (IN REF. 2).
 FT CONFLICT 668 668 E -> D (IN REF. 2).
 SQ SEQUENCE 794 AA; 91799 MW; 74DB3D236224499C CRC64;
 Query Match 100.0%; Score 3104; DB 1; Length 794;
 Best Local Similarity 100.0%; Pred. No. 5.6e-98;
 Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRALSLELMKLNKRNKRETKMRSMVMVQEGMELKLAQTKDLTSKGIKVLQEGKLSIEKE 60
 DB 164 MRALSLELMKLNKRNKRETKMRSMVMVQEGMELKLAQTKDLTSKGIKVLQEGKLSIEKE 223
 QY 61 KIDKCEKTELLEYIQEISCASQDVQVEKCKVDIAQAEEDLKEKDRITLSLKQSLNITFS 120
 DB 224 KIDKCEKTELLEYIQEISCASQDVQVEKCKVDIAQAEEDLKEKDRITLSLKQSLNITFS 283
 QY 121 KOIEDLTVKQQLLETERDNVSKDRERAETLSAEMQILTERLALEROEYKLOQELQSQ 180
 DB 284 KOIEDLTVKQQLLETERDNVSKDRERAETLSAEMQILTERLALEROEYKLOQELQSQ 343
 QY 181 SLQOQEKELSLRLOQLCSFQSEMTSEKNVKEELKLAELDAVQOKEEQSERLVKQLE 240
 DB 344 SLQOQEKELSLRLOQLCSFQSEMTSEKNVKEELKLAELDAVQOKEEQSERLVKQLE 403
 QY 241 EERKSTAEQLTRLDNLLREKEVELEKHHIAHAQAIIAQEKYNDTAQSLRDVTAQLESVQ 300
 DB 404 EERKSTAEQLTRLDNLLREKEVELEKHHIAHAQAIIAQEKYNDTAQSLRDVTAQLESVQ 463
 QY 301 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 360
 DB 464 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 523
 QY 361 SVQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 420
 DB 524 SVQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 583
 QY 421 TAESTNQEYARMVDLQNRSLTKEEIKETSSFLKETTDLKNQLRQDEDFRKLQLEKRG 480
 DB 584 TAESTNQEYARMVDLQNRSLTKEEIKETSSFLKETTDLKNQLRQDEDFRKLQLEKRG 643
 QY 481 KPTAEKENVMTELTWEINKWRLLLYEELYEKTTPFOQQLDAFAEAKOALLNEHGATQEQIN 540
 DB 644 KPTAEKENVMTELTWEINKWRLLLYEELYEKTTPFOQQLDAFAEAKOALLNEHGATQEQIN 703
 QY 541 KIRDSYAQLLGHONLKQKIKHVVKLDNSQLKSVSKLSRVSOLVKRKNELRLOQELDKA 600
 DB 704 KIRDSYAQLLGHONLKQKIKHVVKLDNSQLKSVSKLSRVSOLVKRKNELRLOQELDKA 763
 QY 601 LGIRHFDPSKAFCHASKENFTPLKEGPNCC 631
 DB 764 LGIRHFDPSKAFCHASKENFTPLKEGPNCC 794
 RESULT 2
 HMHR_RAT
 ID HMHR_RAT STANDARD; PRT; 498 AA.
 AC P97779;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Hyaluronan mediated motility receptor (intracellular hyaluronic acid
 DE binding protein) (Receptor for hyaluronan-mediated motility).
 GN HMHR OR IHABP OR RHAMM
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID:10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Vascular smooth muscle;
 RA Savani R.C., Hou G.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.

Best Local Similarity 64.8%; Pred. No. 2.6e-59;									
Matches 411; Conservative 48; Mismatches 87; Indels 88; Gaps 3;									
Qy	1	MRALSLEMLKRNKTRKMRSMVQEGMELKLOATQKDLTSGKIVQLECKLYSIEKE	60						
Db	163	LRLSLLEMLKRNKTRKMRSMVQEGMELKLOATQKDLTSGKIVQLECKLYSIEKE	222						
Qy	61	KIDKCEKTEKLEYIQEISCASDQVEKCKVDIAQLEEDLKEMDKREILSLKQSLSEENIT-F	119						
Db	223	KIDKSETEKLEYIEEISCASDQVEKCKVDIAQLEEDLKEMDKREILSLKQSLSEENIT	282						
Qy	120	SKQIEDLTWKCOLLETERNLVSKDRERAEITLSAEMQILTERLALEROEYKLOQKELOS	179						
Db	283	SKQVEDLVNKCQLLEKEKEDHVRNREHNENLNAEMQNKLQFIEQERREKLOQKELO	342						
Qy	180	OSILOQEKELSLARLOOGLCSFOEEMTSEKNVFEELKALALDAELDAVQKEESERLVKOL	239						
Db	343	DSLLQOKEKELSSSLKQKCSFOEEMVMEKNLFEELKOTLDELKLOQKEQERLVKOL	402						
Qy	240	EEERKSTABQLTRLDNLLREKEVELEKHTIAHAQAAILIAQEKYNDTAQSLRDVTAQLESV	299						
Db	403	EEEAASRAEELKEELKGLKGAELKSAHTQATLLQEKYDSMVQSLSDVT-----	456						
Qy	300	QEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQSL	359						
Db	457	-----	456						
Qy	360	ESVQEKYNDTAQSLRDVSAQLESYKSSLTKEITDLKLENTLOEKVMAEKSVEDVQOOI	419						
Db	457	-----AQFESYKALTASEIEDLKLSENSQERKAAGAKNAEDVQHOI	498						
Qy	420	LTAESTNOEYARMVDQNRSLTKBEETKEITSPFLEKITDLKNLRQODEDFRQLEBK	479						
Db	499	LATESNQEVYRMLDLQTKSALKETEKEITVSLQKITDLQNLKQEEDEFRQLEDE	558						
Qy	480	GKRTAEKNVWTELMWINKWRLLYEELYKYPFQOQLDAFEAEKQALLNHHGATQSQL	539						
Db	559	EGKRAEKENTTAELETEINKWRLLYEELYKYPFQOQLDAFEAEKQALLNHHGAAQSQL	618						
Qy	540	NKIRSYAOLLGHONLKQIKHVVKLKDNLSQKSEVSKLSQVYKRNKONELRLOGELDK	599						
Db	619	NKIRSYAKLGHQNLKQIKHVVKLKDNLSQKSEVSKLSQVYKRNKONELRLOGELDK	678						
Qy	600	ALGIRHFDPSKAFCHASKENF---TPLKEGPNPC	630						
Db	679	VLGIRHFDPSKAFCHASKENFALKTPLEKGNNTNC	712						
RESULT 4									
USOL_YEAST STANDARD; PRT: 1790 AA.									
AC	P25386;								
DT	01-MAY-1992 (Rel. 22, Created)								
DT	01-MAY-1992 (Rel. 22, Last sequence update)								
DE	16-OCT-2001 (Rel. 40, Last annotation update)								
DE	Intracellular protein transport protein USOL.								
GN	USOL OR INT1 OR YDL058W.								
OS	Saccharomyces cerevisiae (Baker's yeast).								
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;								
OC	Saccharomycetales; Saccharomycetaceae; Saccharomyces.								
OX	NCBI_TaxID=4932;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=X2180-1A;								
RX	MEDLINE=91185402; PubMed=2010462;								
RA	Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,								
RA	Yamasaki M.;								
RT	"A cytoskeleton-related gene, usol, is required for intracellular								
RT	protein transport in Saccharomyces cerevisiae.";								
RL	J. Cell Biol. 113:245-260(1991).								
RN	[2]								
RP	SEQUENCE OF 782-1790 FROM N.A.								
RA	Hostettler M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,								

RA	Kendrick K.E.;								
RL	Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.								
RN	[3]								
RP	SEQUENCE OF 1-8 FROM N.A.								
RA	Bai Y., Symington L.S.;								
RL	Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.								
CC	-!- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI COMPLEX.								
CC	-!- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE ER AND THE GOLGI COMPLEX.								
CC	-!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.								
CC	-!- SIMILARITY: BELONGS TO THE VDP/USOL/YBL047C FAMILY.								
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).								
CC	EMBL; X54378; CAA38253.1; -								
DR	EMBL; L03188; AAB00143.1; -								
DR	EMBL; U53668; AAB66659.1; -								
DR	PIR; A38455; A38455.								
DR	SGD; S0002216; USOL.								
DR	InterPro; IP0002017; Spectrin.								
KW	Transp.; Protein transport; Golgi stack; Cytoskeleton; Coiled coil.								
FT	DOMAIN 1 724								
FT	DOMAIN 725 1790								
FT	DOMAIN 465 487								
FT	DOMAIN 991 1790								
FT	DOMAIN 1172 1786								
FT	CONFLICT 847 847								
FT	CONFLICT 924 924								
FT	CONFLICT 1253 1253								
FT	CONFLICT 1319 1319								
FT	CONFLICT 1461 1461								
FT	CONFLICT 1581 1581								
FT	CONFLICT 1600 1600								
FT	CONFLICT 1661 1661								
FT	CONFLICT 1772 1772								
SQ	SEQUENCE 1790 AA; 206424 MW; 6CE2B216E9FD4818 CRC64;								
Qy	Query Match 13.9%; Score 431.5; DB 1; Length 1790;								
Db	Best Local Similarity 24.2%; Pred. No. 4.1e-08;								
Qy	Matches 179; Conservative 149; Mismatches 236; Indels 177; Gaps 33;								
Db	10 KLRNKRRTKMRSMVQEGMELKQATQKDLTSGKIVQLECKLYSIEKEIKDEKCE 69								
Qy	798 ELKNVRDS-LDEMTQRLDVLTKDKENQTALEYKSTHKQEDSITKLEK----- 846								
Db	70 KLEYIQEISCASDQVEKCKVDIAQLEEDL-----KEKD-----REILSLKQ 111								
Qy	847 GLETITSLQKKAEDGINKMGKDLFALSREMQAVEENCKNLQKDKSNVNHQKETSLSKE 906								
Db	112 SLEENT-----FSQIEDLTWKCOLLETER-----DNLVSKDRERAE 150								
Qy	907 DIAAKITEIKATNENLEEMKTCNNLSKEHISKELVEYKSRFQSHDNVLAKLETKLS 966								
Db	151 LS---AEMQILTERL--ALERQVEY-KLQOKELOSO-SLQOKEKLSARLOOQLCSFOEE 203								
Qy	967 LANNYKMQAENESLIKAVEKSNESLIQNLQNKIDMSQOEK-----NFQIE 1016								
Db	204 MTS-EKNVFEELKALAEALDAVQOQ-----EQSERLVKQLEERKSTAEQL 250								
Qy	1017 RGSIEKNI--EQLKKTISDLQTEETISKSDSSKDEYESQISLLKLEKLETATTANDEN 1074								
Db	251 TRLDNLLREKEVELEKHTIAHAQAAILIAQEKYNDTAQSLRDVTAQLESVOE-----KY 303								


```
Db 1353 EQEAKAEQLQRMGSKANSEVAQWRTKYETDAIQTTELEAKKKLAQLQDAEEHVEAVN 1412
Qy 520 ----AFAEKQALLNE-----HGATQEQLN 540
Db 1413 SKCASLEKTKQLQNEVEDLMDVRSNACIALDKKORNFQKVLAEWKQKYETQAELE 1472
Qy 541 -----KIRDSYAQLLGH-----ONLKOKI-----KHVVKLKD 567
Db 1473 ASQKESRSLSTELFKVNAYEBSLDHLETKRENKLNQAEISDLTEQIAEGGKHIHELEK 1532
Qy 568 ENSQLKSEVSKLRSOLVRKKQNELRLQGLDQKALGIR 604
Db 1533 VKKQLDHEKSELQSL-----EEAEASLEHEEGKILRIQ 1566

RESULT 6
MYH4_RABIT
ID MYH4_RABIT STANDARD; PRT: 1938 AA.
AC Q28641;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Myosin heavy chain, skeletal muscle, juvenile.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; TISSUE=Skeletal muscle;
RA Maeda K., Hostinova E., Roesch-Kleinhauf A., Schuster H., Gasperik J.,
RA Wittinghofer A.;
RT "Isolation, sequencing of myosin heavy chain cDNA from rabbit
RT skeletal muscle and a novel cosynthesis of S-1 fragment with the
RT essential and regulatory light chains.";
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MUSCLE CONTRACTION.
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U32574; AAA74199.1; -.
CC HSSP; P13538; 2MYS.
CC InterPro; IPR000048; IQ_region.
CC InterPro; IPR004009; Myosin_N.
CC InterPro; IPR002928; Myosin_tail.
CC InterPro; IPR001609; myosin_head.
CC Pfam; PF00612; myosin_head; 1.
CC Pfam; PF00612; IQ; 2.
CC Pfam; PF01576; Myosin_tail; 1.
CC Pfam; PF02736; Myosin_N; 1.
CC PRINTS; PR00193; MYOSINHEAVY.
CC ProDom; PD000355; myosin_head; 1.
CC SMART; SM00015; IQ; 1.
CC SMART; SM00242; MYSC; 1.
CC SMART;
```

```
DR PROSITE; PS0096; IQ; 1.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
KW Multigene family.
FT DOMAIN 1 783 MYOSIN HEAD-LIKE.
FT DOMAIN 784 813 IQ.
FT DOMAIN 842 1938 COILED COIL (POTENTIAL).
FT NP_BIND 179 186 ATP (POTENTIAL).
FT DOMAIN 658 680 ACTIN-BINDING (BY SIMILARITY).
FT DOMAIN 760 774 ACTIN-BINDING (BY SIMILARITY).
FT MOD_RES 35 35 METHYLATION (MONO-) (BY SIMILARITY).
FT MOD_RES 130 130 METHYLATION (TRI-) (BY SIMILARITY).
FT MOD_RES 552 552 METHYLATION (TRI-) (BY SIMILARITY).
FT MOD_RES 756 756 METHYLATION (SH-1) (BY SIMILARITY).
FT MOD_RES 698 698 ALKYLATION (SH-2) (BY SIMILARITY).
FT MOD_RES 708 708 ALKYLATION (SH-2) (BY SIMILARITY).
SQ SEQUENCE 1938 AA; 223064 MW; D8A8A2EC5B182626 CRC64;

Query Match 13.6%; Score 421; DB 1; Length 1938;
Best Local Similarity 23.0%; Pred. No. 9.9e-08;
Matches 170; Conservative 150; Mismatches 251; Indels 168; Gaps 25;

Qy 9 MKLNRKRETKMRSMVQOE--GMELKLAQTQKDLTESGKIVQLGKLVSTIEKEKIDEKC 66
Db 833 MKLYFKIKPLLSKSAETEKEMANKKEFEKTKESLAKAKKEKELEKVMALQKNDLQL 892
Qy 67 ETEKLLEYIQEISCASDQVEKCV-----DIAQLEEDLKEKREILSLKQSLSEENIT- 118
Db 893 QVQAEADSLADAEERCQDLIKTKOLEAKIKEVTERAEDEEEINNELTAKRKLEDECE 952
Qy 119 FSKQIEDTVVCKQLLETERDNLVSKORERATLSAEMOILTERLALERQEKYEQKELQ 178
Db 953 LKDKIDDELTLAKVEKHKATENK---VKNLTEEMAGLDETIAKTKEKKALQEAHQ 1008
Qy 179 SQSLLOQEKE-----LSARLQOOLCSFQEMTSEKNV-----FKEELKAL-A 220
Db 1009 TLDDLQAEEDKVNTLTAKTKLEQVDDLEGSLEQEKIRMDLERAKRKLEGGDLKLAES 1068
Qy 221 ELDAVQKQEEQSERLVQ-----LE---EERKSTAEQTLRLDNLLEKEVELEKHHAAH 271
Db 1069 TMDIENDKQOLDLKLKKEFMSNLQSKIEDQALAMOLKKEKELQARIELEEEIEAE 1128
Qy 272 QAAILIAQEKYNDPAQSLRDVTAQLESVQEKYNDPAQSLRDVTAQLESQEKYNDTAQSL 331
Db 1129 RASRAKAEKQSDLSRELEISERLE-----EAGGATSAQIENMKKREAEFQKMR 1178
Qy 332 RDVTAQLESQEKYNDPAQSLRDVTAQLESVQEKYNDPAQSLRDVTAQLESVKSTLKEI 391
Db 1179 RD----LEEATLQHEATAATLKKHA--DSVAE-LGEQIDNLQVRVKLEKESLKKWEI 1231
Qy 392 EDLKLLENLTQEKVAMAEKSVEDVQOQILTAESTNQEYARMVQDQL-NRSTLKEE----- 445
Db 1232 DLLASNMTVSKAGNLEKCMKRTLEDQVSELKTRKEEHQRLINDLSAQARLARLTSGEFS 1291
Qy 446 -----EIKETSSFLKTTDLKNQLRQ-----DED-FRKQLE 477
Db 1292 RQLEKDSLVSQSLRGKQAFTQQIEELKRLQEEETKAKSALAHALQSAHDCDLLREQYE 1351
Qy 478 EKGKRTAEKENVMTLMTINKNRLLY-----EELYKTKPQQQLD----- 519
Db 1352 EEQAKAEQLQRMGSKANSEVAQWRTKYETDAIQTTELEAKKKLAQLQDAEEHVEAVN 1411
Qy 520 ----AFAEKQALLNE-----HGATQEQLN-----KIRDSYAQLL 550
Db 1412 AKCASLEKTKQLQNEVEDLMDVRSNACIALDKKORNFQKVLAEWKQKYETQAELE 1471
Qy 551 GHQNLKQ-----KIKHVVK-----LKDENSQLKSEVSKLRSQ----- 583
Db 1472 ASQKESRSLSTELFKVNAYEBSLDQLETLKRENKLNQAEISDLTEQIAEGGKRIHELEK 1531
Qy 584 VKR--KONELRLQGLDKA 600
Db 1532 VKKQVEQKSELQAALAEA 1550
```

```
RESULT 7
MYH2_HUMAN STANDARD; PRT; 1941 AA.
AC Q9UKX2; Q16229; Q14322;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, skeletal muscle, adult 2 (Myosin heavy chain IIA)
DE (MYC-IIa).
GN MYH2 OR MYHSA2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=99318869; PubMed=10388558;
RA Weiss A., Schiaffino S., Leinwand L.A.;
RT "Comparative sequence analysis of the complete human sarcomeric myosin
heavy chain family: implications for functional diversity.";
RL J. Mol. Biol. 290:61-75(1999).
RN [2]
RP SEQUENCE OF 1711-1941 FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=95109625; PubMed=7545970;
RA Smerdu V., Karsch-Mizrachi I., Campione M., Leinwand L.,
RA Schiaffino S.;
RT "Type IIX myosin heavy chain transcripts are expressed in type IIB
fibers of human skeletal muscle.";
RL Am. J. Physiol. 267:C1723-C1728(1994).
RN [3]
RP SEQUENCE OF 1823-1941 FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=95270723; PubMed=7751403;
RA Ennio S., Sant'ana Pereira J., Sargeant T., Young A., Goldspink G.;
RT "Characterization of human skeletal muscle fibres according to the
myosin heavy chains they express.";
RL J. Muscle Res. Cell Motil. 16:35-43(1995).
CC -!- FUNCTION: MUSCLE CONTRACTION.
CC -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
SUBFRAGMENT (S2).
CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF111784; AAD29950.1; -.
DR EMBL; S73840; AAC13916.1; -.
DR EMBL; Z32858; CAA83687.1; -.
DR HSSP; P13538; 2MYS.
DR GeneW; HGNC:7572; MYH2.
DR MIM; 160740; -.
DR InterPro; IPR000048; IQ_region.
```

```
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR InterPro; IPR001609; myosin_head.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF01576; Myosin_tail; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS00096; IQ; 1.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
KW Multigene family.
FT DOMAIN 1 786 MYOSIN HEAD-LIKE.
FT DOMAIN 787 816 IQ.
FT DOMAIN 845 1941 COILED COIL (POTENTIAL).
FT NP_BIND 179 186 ATP (POTENTIAL).
FT CONFLICT 1844 1844 K -> R (IN REF. 3).
SQ SEQUENCE 1941 AA; 681866F83AEA83F CRC64;
Query Match 13.5%; Score 417.5; DB 1; Length 1941;
Best Local Similarity 22.6%; Pred. No. 1.3e-07;
Matches 168; Conservative 145; Mismatches 258; Indels 171; Gaps 23;
Qy 4 LSELMLKLRNKRRTKRMVMYKQEGMELKIQATQKDLTESKGIIVOLEKGLYSIEKEID 63
Db 838 LFFKIKPLKSAETEKEMATKKEEFQKIK-----DELAKESEAKRKELEKMWTLKEKND 892
Qy 64 EKCEPEKLLLEYIQEISCASDQVEKCKV-----DIAOLEEDLKEKREILSLKOSLEEN 116
Db 893 LQLQVQAEAGLDAEERCQDLIKTKIQLEAKIKVETRAEDEEEINAEELTAKRKLEDE 952
Qy 117 IT-FSKQIEDTVKCOLLETEDRNILVSKDRERAETLSAEMLTTERLALEROEYKLOK 175
Db 953 CSELKDDIDDLTLTLAKVEKEKHATENK----VKNLTEEMAGLDETIATKTKKALQEA 1008
Qy 176 ELOSQSLQOKE-----LSARLQOQLCSFOEWTSEKNV-----FKELKIA 218
Db 1009 HQQTLDLQAEEDKVNTITKAKIKLEQQVDDLEGSLEQEKLRMLERAKRKLEGDKIA 1068
Qy 219 LAELDAVQOKEEQSERLVKQLE-----EERKSTAEQLTRLDNLLREKEVELEKHI 268
Db 1069 QESIMDIENEKQDLKKEFEISNLQSKTEDEQALGLOKKIKELQARIEELEESI 1128
Qy 269 AAHAQAIIIAQEKYNDVTAQSLRDVTAQLESVOEKYNDVTAQSLRDVTAQLESQEKYNDTA 328
Db 1129 FAERASRAKAEKQSRDLSRELEFSEIRLE-----EAGGATSAQIEMNKKREAEFQ 1178
Qy 329 QSLRDVTAQLESQEKYNDVTAQSLRDVTAQLESVOEKYNDVTAQSLRDVSAQLESYKSSYL 388
Db 1179 KMRD-----LEATLQHEATATLRKKHA--DSVAE-LGEQIDNLRQVKLEKSEMK 1231
Qy 389 KEIEDKLLENLTLOEKVMAEKSVEDVQOQILTAETNQEYARMVQDL-QNRSTLKEE-- 445
Db 1232 MEIDDLASNVETVSKAGNLEKMCRTLEDQLSELKSEEQORLINDLTAQGRQLQTESG 1291
Qy 446 -----EIKETSSFLEKIDTLKNQLRQO-----DED-FRK 474
Db 1292 EFSRLDEKEALVSLRGKQQAFTQOIEELKRLQEEIKAKNALAHALQSSRHDCDLLRE 1351
Qy 475 QLEEKGRKTAENVMYELTMEINKRWLLY-----EELYKTKPFOOQLDA----- 520
Db 1352 QYEEQESKAELOKALSALSKANTEVAQWRKYETDARTQTEELEAKKLAQRLOAAEEHVE 1411
Qy 521 -----FEAKQALLNE-----HGATQEQELN-----KIRDSYA 547
Db 1412 AVNAKASILEKTKQRLQNEVEDMLDVERTNAAALDKQKRNFDKILAEWKQKCEETHA 1471
Qy 548 QL-----LGHQNLK-----QKIKHVVKLKDENSOLKSEVSKLSQL----- 583
Db 1472 ELEASQKEARSLGTFLFKIKNAYEESLDQLETLRKNKNLQOQEIISDLTQEAEGKRIHE 1531
```


QY 443 KEE-----EIKETTSFLEKTDLNKLRQ-----DE 470
 Db 1285 QSEGEYSQLDEKDTVLSGRQAFTQTEELKRLQEEBKAKSALAHALQSSRHDC 1344
 QY 471 D-FRQLEBKGRFAEKVNTLMEINKWRLLY-----ELYEKTRPFQOOLD-- 519
 Db 1345 DLLREQYEEQAKAELQRAKSKANSEVAQWRKYETDAIQRTTELEAKKLAQLQDA 1404
 QY 520 -----AFPAEQOALINE-----HGATQQLN-----KI 542
 Db 1405 EHVAVNAKASLEKTKQRLQNEVEDLMIDVERTNACAALDKKQRFDKILAEWKQK 1464
 QY 543 RDSYAQLGHQNLKQ-----KIKHVYK-----LKDENSQKSEVSKLRSQ----- 583
 Db 1465 EETHAELEASQKESRSLSTELFKIKNAVEESLDQLETKRENKNIQEQEISDLTEQIABGG 1524
 QY 584 -----VKR--KQNELRLQGLDKA 600
 Db 1525 KRIHELEKIKKQVEQSEKSLQAALSEA 1551
 RESULT 9
 MYH3_RAT
 ID MYH3_RAT STANDARD: PRT: 1940 AA.
 AC P12847;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin heavy chain, fast skeletal muscle, embryonic.
 GN MYH3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87060988; PubMed=3783701;
 RA Strehler E.E., Strehler-Page M.-A., Perriard J.C., Periasamy M.,
 RA "Complete nucleotide and encoded amino acid sequence of a mammalian
 RT myosin heavy chain gene. Evidence against intron-dependent evolution
 RT of the rod.";
 RL J. Mol. Biol. 190:291-317(1986).
 CC -!- FUNCTION: MUSCLE CONTRACTION.
 CC -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 CC SUBFRAGMENT (S2).
 CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X04267; CAA27817.1; -.
 CC PIR: A24922; A24922.
 CC HSSP: P13538; 2MY5.
 CC InterPro: IPR000048; IQ_region.
 CC InterPro: IPR004009; Myosin_N.

DR InterPro: IPR002928; Myosin_tail.
 DR InterPro: IPR002017; Spectrin.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00063; myosin_head; 1.
 DR Pfam: PF00612; IQ; 2.
 DR Pfam: PF01576; Myosin_tail; 1.
 DR Pfam: PF02736; Myosin_N; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR PRODOM: PD000355; myosin_head; 1.
 DR SMART: SM00015; IQ; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS50096; IQ; 1.
 DR Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
 KW Multigene family.
 FT DOMAIN 1 781 MYOSIN HEAD-LIKE.
 FT DOMAIN 782 811 IQ.
 FT DOMAIN 840 1933 COILED COIL (POTENTIAL).
 FT NP_BIND 179 186 ATP (POTENTIAL).
 FT DOMAIN 656 678 ACTIN-BINDING.
 FT DOMAIN 758 772 ACTIN-BINDING.
 FT MOD_RES 130 130 METHYLATION (POTENTIAL).
 FT MOD_RES 696 696 ALKYLATION (SH-1).
 FT MOD_RES 706 706 ALKYLATION (SH-2).
 SQ SEQUENCE 1940 AA; 223857 MW; B5D546A596E5A696 CRC64;
 Query Match 13.4%; Score 415; DB 1; Length 1940;
 Best Local Similarity 22.0%; Pred. No. 1.6e-07;
 Matches 165; Conservative 144; Mismatches 24; Indels 198; Gaps 22;
 QY 4 LSLEMLKRLNRKRETKMRSMVQEGMELKQLQATQDLTESKGIQVQLEKLVLSIEKEKD 63
 Db 833 LFFFKLPLKLSAETEKEMATKEE-----FQTKDELAKSEAKRKELEKLVTLVQEKND 887
 QY 64 -----EKCTEKLEIYQIETSCASDVQEKV-----DIALEEDLKKEDREILSLKOS 112
 Db 888 LQLOVQAENSELLD--AEERC--DOLIRAKFQLEAKIKVEKTEAEDEEINAELETAKRK 943
 QY 113 LEENIT-FSKQLEDLTVKQLLETERDNLVSKDRERAEATLSAEMOITLTERLALEQVEYK 171
 Db 944 LEDECELSKDDIDLELTAKVEKEKHATENK----VKNLTELAGLDETIAKLTRKKA 999
 QY 172 LQOKELQSLSLQBE-----LSARLQQLCSFQEEEMTSEKNV-----FKBE 214
 Db 1000 LQEAHQTLDDLQAEEDKVNLSKLKLEQVDDLESSLEQEKLRVDLERNKRKLEGD 1059
 QY 215 LKALAE-LDAVQKKEQESERLVK-----QLE-----EERKSTAEQLTRLDNLLREKEVL 264
 Db 1060 LKLAQESILDLENDKQDLERLKKDKDFEYSQLQSKVEDQETLSLQLOKKIKELQARIEEL 1119
 QY 265 EKHIAHAQAIIIAQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESQEK-K 323
 Db 1120 EEEIABERATRAKTRKQSDYARELEESERLEE-----AGGVTSQIEUNKKREAE 1171
 QY 324 YNDTAQSLRDVTAQLESE-----QEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVSAQ 379
 Db 1172 FLKLRDLEEAATLQHEATVATLRKKHADSAAELAQIDNLRVKK-----1217
 QY 380 LESYKSTLTKETEDLKNLETLQEKVMAEKSVEDVQOQILTAESTNQYARMVDQLNR 439
 Db 1218 LEKESEFKLEIDDLSSSVESVSKANLEKICRTLEDQLSSEARGKNEETQRSLSLTQ 1277
 QY 440 STLKEEIKEIT-----SSFEKXITDLKNLRQOODE-----471
 Db 1278 KSRLOTEAGELSRQLEEKESIVSQLSRKQAFQIQIEELKRLQLEENKAKNALAHALQSS 1337
 QY 472 -----FRKQLEKGRKRAEKVNTLMEINKWRLLY-----ELYEKTRPFQOQ 517
 Db 1338 RHDCDLLREQYEEQEGRAELQALSKANSEVAQWRKYETDAIQRTTELEAKKLAQR 1397
 QY 518 L-----518
 Db 1398 LQDSEEQVEAVNAKASLEKTKQRLQGEVEDLMVDVVERANSALAAALDKQRNFDKVLAEW 1457


```
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF00612; IQ; 2.
DR Pfam: PF01576; Myosin_tail; 1.
DR Pfam: PF02736; Myosin_N; 1.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00242; MYSC; 1.
DR PROSITE: PS00096; IQ; 1.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW ATP-binding; Methylation; Alkylation; Phosphorylation; Acetylation;
KW Calmodulin-binding; Multigene family; 3D-structure.
FT INIT_MET 0 0
FT DOMAIN 1 782 MYOSIN HEAD-LIKE.
FT DOMAIN 783 812 IQ.
FT DOMAIN 838 840 HINGE.
FT DOMAIN 841 1938 COILED COIL (POTENTIAL).
FT NP_BIND 179 186 ATP (POTENTIAL).
FT DOMAIN 657 679 ACTIN-BINDING.
FT DOMAIN 759 773 ACTIN-BINDING.
FT MOD_RES 1 1 ACETYLATION.
FT MOD_RES 35 35 METHYLATION (MONO-).
FT MOD_RES 130 130 METHYLATION (TRI-).
FT MOD_RES 551 551 METHYLATION (TRI-).
FT MOD_RES 755 755 METHYLATION (MONO-).
FT MOD_RES 697 697 ALKYLATION (SH-1).
FT MOD_RES 707 707 ALKYLATION (SH-2).
FT CONFLICT 907 907 C -> Q (IN REF. 7 AND 8).
FT CONFLICT 980 980 L -> F (IN REF. 1).
FT CONFLICT 1343 1343 E -> D (IN REF. 5).
FT CONFLICT 1545 1545 S -> A (IN REF. 5).
FT CONFLICT 1796 1797 HV -> QL (IN REF. 5).
FT CONFLICT 1830 1830 S -> A (IN REF. 5).
FT CONFLICT 1863 1863 I -> V (IN REF. 10).
FT CONFLICT 1929 1931 IHG -> FH (IN REF. 10).
SQ SEQUENCE 1938 AA; 223013 MW; EDDOICEA2681E10F CRC64;

Query Match 13.3%; Score 413; DB 1; Length 1938;
Best Local Similarity 22.6%; Pred. No. 1.8e-07;
Matches 171; Conservative 133; Mismatches 249; Indels 204; Gaps 24;

Qy 9 MKLNRKRTKMRSMVYKQEG--GMELKLGATQKDLTSGKGVQLVQEGKLVSTKEKIDKDC 66
Db 832 MKLFKIKPLPKLSAESEKEMANMKKEFEKTEELAKSAKEELFEKRMVLLQEKNDLQL 891
Qy 67 ETEKLELYIQETSCASDOVKCKV-----DIAQLEDLKEKREILSLKQSLNENT- 118
Db 892 QVQAEADSLADAEKRCDLIKTKIOLEAKIKEVTERADEEIEINAEULTAKKLEDECE 951
Qy 119 FSKQIEDLVKQCLLETERDNLVSKDRERATLSAEMOILTERLALEREQYEKLOQKELQ 178
Db 952 LKKDIDDLLETLAKVEKEKHATENK-----VKNLTTEMAVLDETIAKLTKEKALQEAHQ 1007
Qy 179 SQSLLOQKE-----LSARLQOQLCSFOEMTSEKNV----- 210
Db 1008 TLDDQLQVEEDKVNLTAKTKLEQVDDLEGSLEQEKRLMDLERAKRKLKGLDLKLAHDS 1067
Qy 211 -----FKEELKALAEALDAVQOKEQSERLVKOLEERKSTABQTLRDLNLLREK 260
Db 1068 IMDLENDQQLDEKLKKDDFELSIOQSKIEDEQALGMQLKKIK---ELQRIEEL--EE 1122
Qy 261 EYELEKHHAAHAQAIIIAQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESE 320
Db 1123 ETEAERTSKRA-----EKHR-----ADLSRELEISERLEPAGGA---TAAQIEMN 1166
Qy 321 QEKYNDTAQSLRDVTAQLESQEKYNDTAQSLR-----DVTQALESVQEKYNDTAQSLRDV 376
Db 1167 KKREAFQKMRD-----LEEAATLQHEATAAALRRKHADSTAE-----GEQIDNLQV 1215
Qy 377 SAQLESYSKSTLKETDLKLENLTQEKVANAQK---SVEDVQOQILTAESTNQEIYARWV 433
Db 1216 KQLEKEKSELKMEIDDLASNMESVSKAKANLEKRCMTLQDLQSEIKTKPEQNG---RMI 1272
Qy 434 QDLQNRSTLKEEIKEIT-----SSPLEKITDLKNQL----- 465
```

```
Db 1273 NDLNTRARLQTTETGEYSRQAEKDALISQLSRGQGFQTOQTEELKRLHEEIKAKNALA 1332
Qy 466 -----ROODEDFRQLEEKGRKTAENKVNMTLMTINKWRLLY-----EELYKT 511
Db 1333 HALQSARHDCLELREQYEEOQAKGELQALRSKANSEVAQWRKYETDAIORTLELEBAK 1392
Qy 512 KPFOQQLD-----AFEAQKQALLNE-----HGATQEOQLNKTRDSYA 547
Db 1393 KKLQRLQDAEHEVAVNAKCALEKTKORLQNEVEDLMDVVERNSNAACAALDKKQKQNF 1452
Qy 548 QLLGHQNLK-----VKR--KQNELRLQELGDKA 600
Db 1453 KILAENKQKYEETOTELEASOKESRSLSTELFKMKNAVEESLDHLETLKRNKNLQOEITA 1512
Qy 578 KIRSQ-----VKR--KQNELRLQELGDKA 600
Db 1513 DLTEQIABEGKAVHELEKVKKHVQOEKSELQASLEEA 1549

RESULT 11
YL17_CABEL STANDARD; PRT; 1130 AA.
ID YL17_CABEL
AC Q11102;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 131.5 kDa protein C02F12.7 in chromosome X.
GN C02F12.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Miller N.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: WEAK, TO MYOSINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U41545; AAK39135.1; -.
DR WormPep; C02F12.7; CE03901.
KW Hypothetical protein; Coiled coil.
FT DOMAIN 121 779 COILED COIL (POTENTIAL).
FT DOMAIN 805 1061 COILED COIL (POTENTIAL).
SQ SEQUENCE 1130 AA; 131485 MW; B0FD2EFE3D99FB09 CRC64;

Query Match 13.1%; Score 408; DB 1; Length 1130;
Best Local Similarity 24.3%; Pred. No. 1.6e-07;
Matches 179; Conservative 132; Mismatches 270; Indels 156; Gaps 32;

Qy 7 ELMKLRNKRRTKMRSMVYKQEG-----MELKQA--TOKDLTSGKGVQLVQEGKLVSTEK 59
Db 290 ELEALRDHEEALKEALLARQDEFHDLQELQLSKRSREDLVSKNDVTALEKKLHNKEK 349
Qy 60 E-----KIDE-KCETKLELYIQETSCASDOVEKCKVDIAQLEDLKEKOR---EILSL 109
Db 350 EVQTLTKELDQVKTEINDKIRRLTEVTEFAFYRK---KFOQOEELRLKRLTLVVEAA 406
Qy 110 KQSLNENT-ESKQIEDLVKQCLLETERDNLVSKDRERATLSAEMOIL-----TER 161
Db 407 KEKLESVSDLOVEVKALKNKVFELEKERENLQSOSESQTQLQSSQVDALEAVLHSTYKE 466
Qy 162 LALERQEYKLOQKELQSQSLIQ--QEKELSAAR---LQOQLCSFOEMTSEKNVFEKELK 216
```



```

QY 114 BENIT-FSKQIEDLVKCOLLETERDNLVSKDRERAETLSAEMOILTERLALEROEYKL 172
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 945 EDECSLKADIDDLLETLAKVEKEKHATENK-----VKNLTTEMASQDSIAKLTKKKAL 1000
QY 173 QOKELQOSLSLOQKE-----LSARLOQCLCSFQEDMTSEKNV-----FKBEL 215
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1001 QEAHQOTLDDLAQEBDKVNTLTAKTKLEQVDDLEGSLEQEKKLRLMDLERAKRKLGGDL 1060
QY 216 KLALAEADVQKEQESERLVKQLE-----EERKSTAEQLTRLDNLLREKEVELE 265
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1061 KLAQESIMDLENEKQSDSEKIKKKDFEISQLLSKIEDQSGAOLQKIKELQARIEELE 1120
QY 266 KHIAHAQAAILTAQBYNDVTAQSLRDVTAQLESVOEKYNDVTAQSLRDVTAQLESBOEYKN 325
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1121 BEIEAERAARAKVEQRADLSRELEISERLE--EAGGATA-----AQIENKKREA 1170
QY 326 DTAQSLRDVTAQLESBOEKYNDVTAQSLRDVTAQLESVOEKYNDVTAQSLRDVTAQLESYKS 385
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1171 EFOKMRD-----LEESTLQHEATAALR--REQADSAE-LGEQIDNLRVQKQKLEKES 1223
QY 386 STLKEIEDL-----KLENL--TLQEKVA-MAEKSVEDVQ-----QILTRES 424
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1224 EYKMEIDLTSMERAVAKANLEKMKRTLEDQSEIKTKSDENVRQNDMNAQARLQT 1283
QY 425 TNQEYARMVQDLQNRSTLKEEIKET---SSFLEKITDLKNQLRQ-----468
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1284 ENGFSRQLEE-----KEALVSLTRGQAYTQOIEELKRHIEEVKAKNALAHAVQS 1336
QY 469 ---DED-FRKQLEEKGTAKENVMTELTKWIRLLY-----EELYEYTKPQQ 516
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1337 ARHDCDLLREQYEEQEAELQKQMSKANSEVAQWRTKYETDAIQTQTEELEAKKLLAQ 1396
QY 517 OLD-----AFEAQKQALLNE-----HGATQEQLN-----540
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1397 RLQDAEESIEAVNSKASLEKTKQRLQGEVDMITDVERANSLEANLQKQNFDPKVLAE 1456
QY 541 ---KIRSYAQLLGHQ-----NLKQIKHVVKLDKENSQKSEVSKLRSQ 582
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1457 WKQYEEQAELEGAQKARSLSLTFPKMKNSEAEALDHLETKRENKLNQOEISDLTEQ 1516
QY 593 L-----VKKQNEURLQGLDQKALGIRHDPDSK 610
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1517 LGETGKSIHEKAKKTVESEKSE--IQTALFEARGTLEHEESK 1558

RESULT 13
MYH3_CHICK
ID MYH3_CHICK STANDARD; PRT; 1940 AA.
AC P02565;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, fast skeletal muscle, embryonic.
OS MYH3.
GN Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RX MEDLINE=87194881; PubMed=3571266;
RA Molina M.I., Kropp K.E., Gulick J., Robbins J.;
RT "The sequence of an embryonic myosin heavy chain gene and isolation
of its corresponding cDNA."
RL J. Biol. Chem. 262:6478-6488(1987).
RN [2]
RP SEQUENCE OF 1502-1940 FROM N.A.
RX MEDLINE=83161144; PubMed=6833296;
RA Kavinsky C.J., Umeda P.K., Sinha A.M., Elzinga M., Tong S.W., Zak R.,
RA Jakovcic S., Rabinowitz M.;
RT "Cloned mRNA sequences for two types of embryonic myosin heavy chains

```

```

RT from chick skeletal muscle. I. DNA and derived amino acid sequence of
light meromyosin."
RL J. Biol. Chem. 258:5196-5205(1983).
CC -!- FUNCTION: MUSCLE CONTRACTION.
CC -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -!- PTM: TWO CYSTEINE RESIDUES IN THE SI DOMAIN ARE SELECTIVELY
ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
SUBFRAGMENT (S2).
CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00430; CAA23712.1; -.
DR EMBL; J02714; AAA48972.1; -.
DR PIR; A02990; A02990.
DR PIR; A29320; A29320.
DR HSP; P13538; 2MVS.
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR InterPro; IPR001609; myosin_head.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF01576; Myosin_tail; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR PRODOM; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 1.
DR Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
KW Multigene family.
FT DOMAIN 1 785 MYOSIN HEAD-LIKE.
FT DOMAIN 786 815 IQ.
FT DOMAIN 844 1940 COILED COIL (POTENTIAL).
FT NP_BIND 179 186 ATP.
FT DOMAIN 660 682 ACTIN-BINDING.
FT DOMAIN 762 776 ACTIN-BINDING.
FT MOD_RES 130 130 METHYLATION (TRI-) (POTENTIAL).
FT MOD_RES 700 700 ALKYLATION (SH-1).
FT MOD_RES 710 710 ALKYLATION (SH-2).
FT VARIANT 379 379 G -> D.
FT CONFLICT 1547 1547 T -> A (IN REF. 2).
FT CONFLICT 1913 1915 ERA -> GRT (IN REF. 2).
SQ SEQUENCE 1940 AA; 222816 MW; C34833D75B04DF2 CRC64;

Query Match 13.0%; Score 404; DB 1; Length 1940;
Best Local Similarity 23.2%; Pred. No. 3.7e-07;
Matches 177; Conservative 138; Mismatches 233; Indels 214; Gaps 29;

Qy 9 MKLRNKRKTFMRSMMVKOE--GMELQATOKDITFESKGIQVLEKGLVSIETKEIDKDC 66
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 835 MKLFKFKPLPLKASAESEKEMANNKEEFKTEELAKSAKKEELKMWVSLQKRNQLQ 894
Qy 67 ETEKLEVIQETSCADQVEKCKV-----DIAQLEDLKEKDEILSLKQSLSEENIT- 118
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Db 895 QVQAEADGLADEERCDQLIKTKIQLEAKIKELTERAEDEEEMNAELTAKKRKLEDCSE 954
QY 119 FSKOIEDLVKCOLLTERDNLVSKDRERAETLSAEMOILTERLALEROEYKLOQKEIQ 178
Db 955 LKKDIDDLETLAKVEKEKHATENK-----VKNLTEMAGLDETIAKTKKALQEAHQ 1010
QY 179 S-----QSLQOQEKEL-----SARLQOQLCFSQB- 202
Db 1011 TLDDLQAEEDKVNLTAKTKLEQQVDDLEGSLEQEKLRMDLRAKRLKGLDKTKWTQS 1070
QY 203 --PWTSEKVNFKELKALAEALDAVQOKEBQSERLVKQLEBEERKSTAEQTLRLDNLRLRK 260
Db 1071 TMDLENDKQOLDEKLKKDFEISQSKIEDEQALGMQLOKKIK---ELQARIEEL--BE 1125
QY 261 EVELEKHIAHAQAIIIAOEKYN-DFAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLES 319
Db 1126 EIAEETSRAKA-----EKHRADLSRELEISERLE---EAGGATAQI-DMNKKREA 1174
QY 320 EQEKYNDTAQSLRDVTAQLESE-----QEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRD 375
Db 1175 EFQKMR---RDLEEATLQHEATAAALRKHKH---ADSTADVGEQIDNLQR----- 1217
QY 376 VSAQLESYSKSTLKETEDLKENLTLQEKVMAEK---SVEDVQOQILTAESTNQ----- 427
Db 1218 VKQLEKSEKSELKMEIDDLASNMESYSKAKANLCRSLQSLSEIKTKKEEQOQTIND 1277
QY 428 -----EYARMYQDLQNRSTLKEEIKETT---SSFLEKITDLKNLQROQ--- 468
Db 1278 ISAKARLOTSEGEYSRQVEE-----KDALISLSRQKQFTQOIEBLKRLHEEIK 1330
QY 469 -----DED-FRQLEEKGRTRAKENVNTEMTMEINKWRLLY-----BE 506
Db 1331 KKCPAHALQSARHDCDLLREQVEEQEAKGELQALSKANSEVAQWRTKYETDAIQTREE 1390
QY 507 LYEKTFPFOQLD-----AFEAQKQALLNE-----HGATQOLNKKI 542
Db 1391 LEEAKKLLAQRLQDAEEHVEAVNSKQASLEKTQORLQNEVEDLMDIVERSNAACAALDK 1450
QY 543 RDSYAQLLGHQNLK-----KIKHVVKLKDENSOL 572
Db 1451 OKNFDKILSEWKQYBETQAELEASQESRSLSLSTELFKMKNAYESLDHLETLKRENKL 1510
QY 573 KSEVSKLRSQ-----VKR--KONELRLOGLDKA 600
Db 1511 QOEISDLTEQIAEGGKRAIHELEKVKKQIEQEKSELQTALEEA 1552
RESULT 14
MYSS_RABIT STANDARD; PRT; 1084 AA.
AC P02562;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Myosin heavy chain, skeletal muscle (fragments).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE OF 1-258.
RA Capony J.-P., Elzinga M.;
RT "The amino acid sequence of a 34,000 dalton fragment from S-2 of
RT myosin."
RL Biophys. J. 33:148A(1981).
RN [2]
RP SEQUENCE OF 259-428.
RX MEDLINE=85131142; PubMed=3972832;
RA Lu R.C., Wong A.;
RT "The amino acid sequence and stability predictions of the hinge
RT region in myosin subfragment 2".
RL J. Biol. Chem. 260:3456-3461(1985).
RN [3]

RP SEQUENCE OF 409-1084 FROM N.A.
RX MEDLINE=87304245; PubMed=3305014;
RA Maeda K., Sczakiel G., Wittinghofer A.;
RT "Characterization of cDNA coding for the complete light meromyosin
RT portion of a rabbit fast skeletal muscle myosin heavy chain.";
RL Eur. J. Biochem. 167:97-102(1987).
CC -!- FUNCTION: MUSCLE CONTRACTION.
CC -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -!- SIMILARITY: THE PERIODICITIES OF HYDROPHOBIC AND CHARGED RESIDUES,
CC WHICH DICTATE THE ALPHA-HELICAL COILED-COIL STRUCTURE ARE
CC CONSERVED.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; X05958; CAA29391.1; -.
DR PIR; A02985; A02985.
DR PIR; A05280; A05280.
DR PIR; S00084; S00084.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW ATP-binding; Multigene family.
FT NON_TER 1 1 ALPHA-HELICAL TAILPIECE (S2).
FT DOMAIN <1 >258
FT NON_CONS 258 259
FT DOMAIN <259 1084
FT DOMAIN 455 1084
FT VARIANT 405 405
FT VARIANT 408 408 L -> V.
FT VARIANT 421 421 V -> L.
FT VARIANT 423 423 E -> D.
FT VARIANT 423 423 S -> G.
FT VARIANT 426 426 K -> R.
SQ SEQUENCE 1084 AA; 125498 MW; 229CFD69A5E1P7F0 CRC64;
Query Match 12.9%; Score 400.5; DB 1; Length 1084;
Best Local Similarity 22.5%; Pred. No. 2.8e-07;
Matches 166; Conservative 141; Mismatches 254; Indels 177; Gaps 25;
QY 16 ETKRSMVMVQEGMELKLOATQKDLTSGKIVQLEGKLVSTKID-----EKCETELK 71
Db 3 ETEKMANKEE-----FEKTESLAKAEAKRKEEKVLMQEKNDLQVQAEADSL 57
QY 72 --LEYIQISCASQVCKVDIAQLEEDLKEKDEILSLKOSLEENIT-FSKQIEDLV 128
Db 58 ADAERQDLIKTKQLEAKIKETVTERAEDEEINAEALTAKRKLKDECESELKDDIDLEL 117
QY 129 KCOLLETERDNLVSKDRERAETLSAEMOILTERLALEROEYKLOQKESLQSQLEQKE 188
Db 118 TLAKVEKEKHATENK-----VKNLTEMAGLDETIAKTKKALQEAHQQTLDLQAEAD 173
QY 189 -----LSARLQOQLCFSQEMTSEKNV-----FKEELKALAEALDAVQOKEQ 231
Db 174 KVNLTAKTKLEQQVDDLEGSLEQEKIRMDLERAKRKLKGLDKLAQETSHNDIENDKQ 233
QY 232 SERLVKQLE-----EERKSTAEQTLRLDNLRLREKEVELEKHIAHAQAIIIAOEKY 282
Db 234 LDEKLLKLEFFNTNLSQKIEDEQALMTNLRQIEEL--EEIEAER--ASRAKA-----EKQ 284
QY 283 NDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESEQ 342

Query Match	12.98;	Score 400;	DB 1;	Length 1940;
Best Local Similarity	21.98;	Pred. No. 5e-07;		
Matches 168;	Conservative 137;	Mismatches 261;	Indels 200;	Gaps 22;
QY	4	LSLEMLKRNKRETKMRSMVMYKQGMELKLCATQKDTESKGVKIVOLEGKLVSTEKKEKID	63	
DB	833	LFPIKPLKSAETEKEMATWKEE-----FQTKDELAESAFAKKELEBKLVTVQEKND	887	
QY	64	-----EKCETEKLLFYIOEISCAQDVCKV-----DIAQLEEDUKEDKREILSLKQS	112	
DB	888	LQLOVQAESNLLD--ABERC--DOLIKAKFOLEAKIKEVTERAEDDEEINAEILTAKKRK	943	
QY	113	LEENIT-FSKOIEDLTVKCOLLETERRNLVSKDRERAETLSAEMQILTERLALAEQVEK	171	
DB	944	LEDCSELKDDIDLEULTLAKVEKEHATENKVNKLTEELSG----LDETAKLTREKKA	999	
QY	172	LQKRELQSLSLQOEKE-----LSARLQOLCSFOEEMTSEKNV-----PKEE	214	
DB	1000	LQEAHQALDLOAEEDKVNLSNKTSKLEQOVEDLESSLEQEKKLVDLERNKRLG	1059	
QY	215	LKLALAE-LDAVQOKEQSOSERLVK-----OLE---BERKSTABQLTRLDNLLREKEVL	264	
DB	1060	LKLAQESILDLQNDKQDLDERLKKKDFEYCOLQSKVDEQTLGLQFOKKIKELQARTEEL	1119	
QY	265	EKHATAHAQAALLIAQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESEQE-K	323	
DB	1120	EEETEAERATRAKTEQORSYARELEELSELEE-----AGVTSTQTELNKKREAE	1171	
QY	324	YNDTAQSLRDVTAQLESE---QEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVSAQ	379	
DB	1172	FLKLRRDLEATLQHEAMWATLRKKHADSVAELEQIDNLQVRQK-----	1217	
QY	380	LESYKSTLKBIEDLKLNTLQEKVMAEKSVEDVQOQILTAESTNOEYARMYODLQNR	439	
DB	1218	LEKESKFLEKIDDLSSMESVSKANLEKICRTLEDQLESEARGKNEIQRSLSLTQT	1277	
QY	440	STLKEEELKEIT-----SSPLEKITDLKNOLROODEDFRKOLEKGR-----	482	
DB	1278	KSLQTTAGELSLROLEKESIVSOLSKSAFTQOTTELKQLEEEKNKALNAHAHQSS	1337	
QY	483	-----TAKENYMTLMTINKWRLLY-----BELYEKTKPQQQ	517	
DB	1338	RHDCDLLREQVEEPEGKAEQLQALSKANSEVAQWRTKYETDAIQTEELEEAKKLAQR	1397	
QY	518	L-----	518	
DB	1398	LQDSEEOEAVNKAACASLEKTKQKRLQGEVEDLMYDVERANSALAAALDKORNFQKVLAEW	1457	
QY	519	----DAFEAKQKALLNEHGATQEOQLNKIRDSYAQLLG-----HONLKQIKHVVYKLK	566	
DB	1458	KTKCEESQAELEASLKESRSLSLSTFLKNAYEBALQOETVVKRENKLEGEIADLTQEI	1517	
QY	567	DENSQLKSEVSKRSQVLVKKR-----QNELRLOQELDKALGIR	604	
DB	1518	AENKTTHELEKSKQTELEBKADIQALEPAEAALAEHEEAKILRIO	1563	

Search completed: December 20, 2002, 15:12:10
Job time : 24.3068 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:11:28 ; Search time 7.91077 Seconds
(without alignments)
1337.138 Million cell updates/sec

Title: US-09-685-010-48

Perfect score: 3104

Sequence: 1 MRALSLEKLNRKRETKMR.....FCHASKENFTPLKEGNPNCC 631

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 106657 seqs, 16763532 residues

Total number of hits satisfying chosen parameters: 106657

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
1	349	11.2	1958	12 US-10-028-946-4
2	349	11.2	2054	12 US-10-028-946-2
3	347	11.2	868	9 US-09-884-001-19
4	345	11.1	2053	9 US-10-017-216-2
5	342	11.0	2055	9 US-10-017-216-4
6	338.5	10.9	879	9 US-10-108-605-217
7	331	10.7	1597	9 US-10-017-216-6
8	324	10.4	1641	9 US-10-017-216-5
9	320.5	10.3	2125	10 US-09-919-172-29
10	302.5	9.9	1045	10 US-09-815-242-10617
11	302.5	9.7	2568	10 US-09-866-108-3
12	300	9.7	909	10 US-09-925-299-988
13	299.5	9.6	660	10 US-09-864-761-47959
14	299.5	9.6	677	10 US-09-745-763-168
15	299	9.6	2310	9 US-09-991-496-120
16	299	9.6	2310	10 US-09-874-923-120
17	298	9.6	2139	10 US-09-727-384-6
18	288.5	9.3	689	9 US-10-108-605-305
19	286.5	9.2	1786	9 US-09-742-096-3

20	284.5	9.2	1192	10	US-09-815-242-10903	Sequence 10903, A
21	278	9.0	751	10	US-09-864-761-38419	Sequence 38419, A
22	274.5	8.8	900	12	US-10-071-751-21	Sequence 21, Appl
23	273.5	8.8	1711	10	US-09-771-161A-219	Sequence 219, App
24	273.5	8.8	1711	10	US-09-771-161A-220	Sequence 220, App
25	267.5	8.6	2285	10	US-09-932-183A-2	Sequence 2, Appl1
26	266	8.6	1179	10	US-09-815-242-13262	Sequence 13262, A
27	265.5	8.6	945	10	US-09-745-763-191	Sequence 191, App
28	265	8.5	830	9	US-10-033-245-7	Sequence 7, Appli
29	265	8.5	830	9	US-10-033-223-7	Sequence 7, Appli
30	265	8.5	830	9	US-10-033-167-7	Sequence 7, Appli
31	265	8.5	830	12	US-10-033-246-7	Sequence 7, Appli
32	265	8.5	830	12	US-10-033-301-7	Sequence 7, Appli
33	265	8.5	830	12	US-10-033-326-7	Sequence 7, Appli
34	265	8.5	1286	9	US-10-017-216-7	Sequence 7, Appli
35	263.5	8.5	600	10	US-09-975-901-2	Sequence 2, Appli
36	258	8.3	1184	10	US-09-815-242-5229	Sequence 5229, Ap
37	258	8.3	1188	10	US-09-815-242-12125	Sequence 12125, A
38	258	8.3	1242	10	US-09-925-299-911	Sequence 911, App
39	255.5	8.2	981	10	US-09-815-242-12211	Sequence 12211, A
40	255	8.2	1179	10	US-09-815-242-13608	Sequence 13608, A
41	244.5	7.9	1048	10	US-09-741-669-409	Sequence 409, App
42	244.5	7.9	1048	10	US-09-815-242-10062	Sequence 10062, A
43	244	7.9	435	10	US-09-866-582-33	Sequence 33, Appl
44	241.5	7.8	2835	10	US-09-885-535-4	Sequence 4, Appli
45	241	7.8	996	10	US-09-815-242-5251	Sequence 5251, Ap

ALIGNMENTS

RESULT 1
US-10-028-946-4
; Sequence 4, Application US/10028946
; Patent No. US20020123622A1
; GENERAL INFORMATION:
; APPLICANT: Yu, Xuanchuan
; APPLICANT: Miranda, Carl Johan
; TITLE OF INVENTION: No. US20020123622A1el Human Kinases and Polynucleotides Encodi
; FILE REFERENCE: LEX-0289-USA
; CURRENT APPLICATION NUMBER: US/10/028-946
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/258,335
; PRIOR FILING DATE: 2000-12-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1958
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-028-946-4

Query Match 11.2%; Score 349; DB 12; Length 1958;
Best Local Similarity 22.3%; Pred. No. 6.3e-09;
Matches 154; Conservative 151; Mismatches 239; Indels 148; Gaps 29;
Qy 25 KQEGNELQATQDLTESKGIKIVOLEGKLYSI-----EKEKIDKEKCEK-L 71
Db 443 KTSSEMKKLLIKSKELQSDQCKHKEQEMTLRRRVSEVAVLSQKEVELKASQKSL 502
Qy 72 LE-----YIQETSCASDOVEKCKVDIA-----OLEEDLKEKDEILSLKOSLEENITF 119
Db 503 LEQDLATVITECCSLKRSLEQARMEVSQEDDKALQLLHDIREQSKLQEIKEQ-----EY 557
Qy 120 SKQIEDLVKCOLLETEDNLSKDRERAETLSAEMQILTERLALER----- 166
Db 558 QAQVEEMRLMNQLE---EDLVA-RRRSLDYSELRL--ESRLAAEFKRRKATECOHKL 611
Qy 167 -----QYEYKLOKEQLQSLSLQKELSLRLOQLQCSFOEFTSEKNVFEKELK 216
Db 612 KAKDQKGPVEGYAKLEKINAEQQLKIQ---ELQEKLEKAVKASTEATELLQNI--ROAKE 667


```
Db 121 SVI--EVTKGQLEVIQTVTAKEVIOGEVRCCLKLELDTESQAE--QERDAAAQLQA 176
Qy 166 EKELSRARQQQLCSQFEEMT-----SEKNVFEELKALAEALD-----AVQOKEQS 232
Db 177 BOEGTALQOQAAHEKENVOLREKWEKERSWHOQELAKALESERKMELEMRUKEQOT 236
Qy 233 ERLVKQLEERKSTAEQULTRLDNLLREKEVELEKHIAHAQAAILIAQEKYNDTAQSLRDV 292
Db 237 EMEAIAQAE-----EERTQAESALCQMLETERKERSVLSLETTLTQKELADASQOLERL 291
Qy 293 TAQLESVOEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESE-----QEKYND 347
Db 292 RQDMVKQIKKEQETTILQ---TQLEAQARELKEAARQHRDDLAALQBESSSLQDKM-D 347
Qy 348 TAQSLRDVTAQ-----ESVOEKYNDTAQSLRDVSAQLESYKSSITLKEIEDLKL 396
Db 348 LQKQVEDLKSQVLAODDSQRLVQEVQEKLRET-QEYNRIQKELERKASLTLSMERKEQ 406
Qy 397 ENLTLOEKVAMAEKSVEDVQOQILTAESTNOEYARMVDLQNRSTLKBEIEKEITSSFL 456
Db 407 RLLVLQEADSIQOELSALRQDMQEAQGEQKELSAQME-----LLRQEVKEKEADFLA 459
Qy 457 KITDLKQNRQODEDFRQLEKEGKRTAEKENVMTELMEINKW--RLLYEELYEKTQPF 514
Db 460 QEAQLEEL-----EASHITEQQLRASLWAQEAQAKAQLQLRLRST 499
Qy 515 QOQLDAFAEKQ-----ALLNEHGATQOLNKRIRDSYAQLLG----- 551
Db 500 ESQLEALAEQPGNQAAQALASLYSALQALGVCSERPELGGGSDSPVWNGLEPFD 559
Qy 552 -----HQNL--KQIKHWV-----KLKD-----E 568
Db 560 QNGARSLFKRGPLLTALSAAVASALHKLHQDLWKTQOTRDVLRDQVKLEBLTDTAE 619
Qy 569 NSQLKSEVSKRSQVLR-----KONELR-----LQSELDKA-LGIRHF 606
Db 620 KSQVHTEQLQDLQRLQSONQEEKSKWEGKQNSLELMELHETMASLQSLRRAELQRM 679
Qy 607 DPSKAFCHASKENFT 621
Db 680 QGERLLQAQENLT 694
```

```
RESULT 4
US-10-017-216-2
; Sequence 2, Application US/10017216
; Patent No. US20020160483A1
; GENERAL INFORMATION:
; APPLICANT: KAPELLER-LIBERMANN, Rosana
; TITLE OF INVENTION: 13245, A No. US20020160483A1el Human Myotonic Dystrophy Type Prot
; FILE REFERENCE: 10147-5701
; CURRENT APPLICATION NUMBER: US/10/017,216
; CURRENT FILING DATE: 2001-10-23
; PRIOR FILING DATE: 2001-10-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2053
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-216-2
```

```
Query Match 11.1%; Score 345; DB 9; Length 2053;
Best Local Similarity 22.4%; Pred. No. 1e-08;
Matches 152; Conservative 154; Mismatches 236; Indels 136; Gaps 29;
```

```
Qy 25 KQEGMELKQATQKDLTSEKQIVOLEGKLVSI-----EKEKIDKCEK-L 71
Db 443 KTSMEKKLLIKSKELQDSQDKCHKMEQEMTLRHRVSEVAVLSQKEVELKASETQRS 502
```

```
Qy 72 LE-----YIQEISCASQDVQEKCKVDIA-----QLEEDLKEKDRILSKOSLEENITF 119
Db 503 LEQDLATYITECSSLKRSLQEARMEVSDDDKALQLLHDIROSRKLOKEQ-----EY 557
Qy 120 SKQIEDLVKCOLLETERDNLVSKORERAEITSAEMQIILTERIALB--ROEYEEKLOOKEL 177
Db 558 QAQVEMRLMNQLE---EDLVA--RRSDLYESEL--ESRLAAEFERKATECOHKLL 611
Qy 178 OSOSLLOB-----KELSARLOOQLCSFOREMTSEKNVFEELKALAEALDAVQOKE 230
Db 612 KAKQDGKPEVGAYAKLEKINAEQQLKIQLEKLE-----KAAKERARELEKLNQRED 665
Qy 231 QSERLVKQI---EERKSTAEQULTRLDNLLREKEVELEKHIAHAQAAILIAQEKYNDTAQ 287
Db 666 SSEGIRKKLVAEERHSLKENVKRLTWER--RENKLDQDIQTKSQI-----Q 713
Qy 288 SURDVTQLESVOEKYNDTAQSLRDVTAQLESQEKY-----NDTAQSLRD---VTA 336
Db 714 QWADKILELE---EKHREAQVSAQHLEVHLKQEQHYEEKIKYVDNQIKKDLADKETLEN 770
Qy 337 QLESQEKYNDTAQSLRDVTAQLESVOEKYNDTAQSLRDVSAQLESYKSSITLKEIEDLKL 396
Db 771 MQRHEEAHAKGKILSEQKAMINAMDSKIRSLQRIVELSEANKLAANSSLFTQRMKA 830
Qy 397 ENLTLOEKVAMAEKSVEDVQOQILTAESTNOEYARMVDLQNRSTLKBEIEKEIT---SS 453
Db 831 Q-----EEMISELRQCKFYLETQAGL-----EAQNRKL--EQLEKISHQDHS 872
Qy 454 FLEKTYDLKNQIR-----QODEFRKQLEEGKRTAEKENVMTEL-----TMEINKWR 501
Db 873 DKNRLLETRLRREVSLSEHEEQKLELQKUTELQLSQERESQTLQAAARALE--SOLR 931
Qy 502 LLYEELYEKTQFOQLDAFAEKQALLNEHGA-----TOEQLANKIRDSYAQLLG 551
Db 932 QAKTELETTAEAEETIQTALTAHRDEIQRFDALRNSCVITDLEQLNQLTEDNAE-LN 990
Qy 552 HQN--LKQIKHWKLKDNSQLKSEVSKRSQVLRQKQNELRQGLDQKALGIRHFDPS 609
Db 991 NQNFVLSKQLDEASGANDIEVLQSEVDHLRREITEREM--QLTSQKQTMEAL----- 1041
Qy 610 KAFCHASKENFTPLKEGN 627
Db 1042 KTTCTMLEEQVMDLEALN 1059
```

RESULT 5

```
US-10-017-216-4
; Sequence 4, Application US/10017216
; Patent No. US20020160483A1
; GENERAL INFORMATION:
; APPLICANT: KAPELLER-LIBERMANN, Rosana
; TITLE OF INVENTION: 13245, A No. US20020160483A1el Human Myotonic Dystrophy Type P
; FILE REFERENCE: 10147-5701
; CURRENT APPLICATION NUMBER: US/10/017,216
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/242,429
; PRIOR FILING DATE: 2000-10-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2055
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-017-216-4
```

```
Query Match 11.0%; Score 342; DB 9; Length 2055;
Best Local Similarity 22.0%; Pred. No. 1.4e-08;
Matches 152; Conservative 152; Mismatches 240; Indels 148; Gaps 29;
Qy 25 KQEGMELKQATQKDLTSEKQIVOLEGKLVSI-----EKEKIDKCEK-L 71
Db 442 KYSSMEKKLLIKSKELQDSQDKCHKMEQEMTLRHRVSEVAVLSQKEVELKASETQRS 501
```



```
Query Match      10.3%; Score 320.5; DB 10; Length 2125;
Best Local Similarity 22.5%; Pred. No. 1.3e-07;
Matches 151; Conservative 141; Mismatches 253; Indels 125; Gaps 28;

Qy 7 ELMKLRNREYKMRSMVMYKQEGMEKMQATQKDLTSGKIVQLEGGKLVSTEKEDK 66
Db 636 ELPVREAAENELRQQRNVEDISL-----QKIRASEAK--QYRELETIVREK--EAA 686
Qy 67 ETEKLEVIQETSCASDVQEKCKVDIAOLEEDLKEKDRILSLKOSLEENITFSKQIEDL 126
Db 687 ERE--LERVQLTIFAERK-----AAVEENL-----INFRQLEENITFRTLEDH 731
Qy 127 TVKCOL-----LETERDNLVS---KDRRAETLSAEMQILTERLALERQYKLEKQKELQS 179
Db 732 LKRKDLNDLEQOKNKLMEELRRKRDNEELLKLKQMEKDLAFQKQVAKK-QLKE--- 787
Qy 180 QSLQOEKELSLAQ-----QQLC-----SFOEMTEKKNVFEEL 215
Db 788 ----KQTELEARRKITEIQTCTRENALPVPITQATSCRAVTGLQEHDKQK-----ABEL 840
Qy 216 KLALAEIDAVQOKEROSER-LVKOL-----EERKSTAEOLTRLDNLLREKEVEL 264
Db 841 KOQVDLTAANKKABQDRELYELNALOLEKTSSEKARLLKDKLDETNTLRLCKLEL 900
Qy 265 EKHAHAHAQAILIAQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLES---EQ 321
Db 901 ER---KQAEKGYSQLRELGRQLNQTTKAAEEAMQASDLKKIKRNVQLELESINHEK 956
Qy 322 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESVQ--EKYNDTAQSLRDVSAQ 379
Db 957 GKLOREVDRIITRAHAFAVKNTQHLNSQIHSFRD-EKELERLQICQKSD-----HLKEQ 1009
Qy 380 LESYKSSYTLKETEDLKLNLTLQ---EKVAMAEKSVEDVQOQILTAESTNOEYARMVDL 436
Db 1010 PEKSEQLLQNKAKKENNDKIQRNELEKSNCEAKMLKQKVELTRQNETKLMQRI 1069
Qy 437 QNRS---TLKEEETKEITSSFLEKTTDLKLNLRQODE-----DFRKOLEKKGKRT 483
Db 1070 QAESENVILEKQTIQORCEALKIQAQDFKQDLRSTNEHLHKQTKTEQDFQRKIKCLEEDL 1129
Qy 484 AKEKNVMTLMEINKWRLLYELYEKTPFOOQLDAFEAK-----QALLNEHGATQE-- 537
Db 1130 AKSNLVSEFKQKQDQOQIITQNTKTRVNLNAELNASKERKRGQKVLQQAQVQELN 1189
Qy 538 -QLNKIRDSYAQLLGHQNLKQKIKH--VVKLDKNSQLKSEVSKLSQLVKRRKQNELRLQ 594
Db 1190 NPLKVQDE---LHLKTEEQMTRKKVLPQESGKFKQSAEERFKMKELMESKVITE 1245
Qy 595 GELDKALGIR 604
Db 1246 NDIS---GIR 1252

RESULT 10
US-09-815-242-10617
; Sequence 10617, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: zyskind, Judith W.
; APPLICANT: wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815.242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10617
; LENGTH: 1045
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10617

Query Match      9.9%; Score 306.5; DB 10; Length 1045;
Best Local Similarity 21.0%; Pred. No. 2.7e-07;
Matches 156; Conservative 148; Mismatches 255; Indels 185; Gaps 28;

Qy 7 ELMKLRN---RRETKMRSMVMYKQ-EGMELKQATQKDLTSGKIVQL-----50
Db 98 ELFKRGGTGVRKQSAKISLIVKDAQGKELRSYKRRVEDLIQELHLNANOFAQIVLIP 157
Qy 51 EGKLVSIKEKIDCEKTEKLEKLEIYQIEISCASDVQEKCKVDIAOLEEDLKEKDRILSLK 110
Db 158 QGEFTFLIAKSDK---EKVLNL-----FGTELYOLFSENJERKLIANTQEIQTQ 207
Qy 111 OSLE---ENITFSKQIE---DLTVKCOLLETERDNLVSK---DREARETLSAEMQ-----156
Db 208 QKIELKTLQHLHSEPEPTMTLFEKLQLELSEQOQAEQALLVERQIATLQAKQAKQEV 267
Qy 157 --ILTERLALERQYKELQO--KELQSLSLQOEKELSAKL-----QQOLCSFQSEMISE 207
Db 268 RYAIERONLQOQKELLEKKAQVEQETVIERLEQIQQLKWSQKQSLAEKVEKRS 327
Qy 208 KNVFKEELKALAEIDAVQOKEEQSERLVKOLEERKSTAEOLTRLDNLLREKEVELEKH 267
Db 328 KQOEKQTSKQOQALMETQOALTQWQAIMSELEQOPLIAEKQERLQTIQRO-----379
Qy 268 IAAHAQAAILIAQEKYNDTAQSLRDVTAQLESQEKYNDTAQ---SLRDVTA-----QLE 318
Db 380 -----LPQYQVEYQLAQOQIAEQANYQAIQKEYESCQOQEKITLADKAVATAKQFIEQ 431
Qy 319 SEQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVSA 378
Db 432 GTLEKANFECSSVADHWQNFVERWQKNQKAWQKISQNVQVELHQLTQFAVEQOQKSAEPA 491
Qy 379 QLESYKSSYTLKEIEDLKLNLTL-----QEKVAMAEKSVEDVQO 417
Db 492 KLQTKKS-----QWASLQIQRUSLLEEGEPCPVCSLSHPKQOQTHQEVSLSEB--IDQAE 545
Qy 418 QILTAESTNQYARMVDL-----QNRSTLKE-----EIKEITSSFLE-----456
Db 546 ELTEVEKTVQRTETLSALGAEKQKESQLOQOEAAAYTEEQQLAAQADLQPLTLGLTF 605
Qy 457 -----KITDLKNLRQODEDFRKOLEE---KGKRTAKENVMTELTMWEINKWR-----501
Db 606 SQVTPAETAEVSQLAKEKQIAQKLTETSVKORLAELEQVAENSORFVLRQOQVETM 665
Qy 502 -----LTYEELYEYKTPQOQQLDAFEAEKQALLNEHGATQEQOL 539
Db 666 QOSLERITIQOQMIASQLLDATVYEMTKQOQALLOEELSAPFQKENVTTQ--GETLKR 724
Qy 540 NKIRDSYAQLLGH-----QNLKQKIKHVVK-----LKDNSQLKSEVSKLSRQ--582
Db 725 EMILES---TLTHLEKEQOTLQOTVAQLESQNLNAVLTQEGVTEQDQLTWELKEVPTLESQ 781
```


Matches 138; Conservative 124; Mismatches 218; Indels 121; Gaps 26;

```
Qy 5 SLEMLKRNKRETKRMVMVK--QEGMELKQATQKDLTESKG-KIVQLEGKLV-----56
Db 90 TVELOQRNHHKDSQFETDIVQRMQEQTKLEQLRAELDEMYGQOIVQMKEQLRQHMAQ 149
Qy 57 IEKEKIDCEKTEKLLIYQIBISCASQVQKCKVDIAQLEBDL-----KEDREIISLK 110
Db 150 MEENKTRHKGEMENALRSYNTVNEQIKLMVAINELNKLQDTNSQKELKEELGL- 208
Qy 111 QSLSENITFSQIIDLTVKCOLLETRDNLVSKDREARLTSAMQILTERALEREYE 170
Db 209 -ILEKCALQRLQEDLV---EELSFSRQI-----QARQTIARQEQ---SKLNEAHKSL 256
Qy 171 KLOQKELQSLLQKQKELSLARLQQLCSFO---EEMTSEKNVFKPEEL-KIALAELDAVQ 226
Db 257 TVEDLKAEIVSASESRKELEKHEAEVNTYKIKLEMLEKEKNVLDRAESQEAELERL 316
Qy 227 QK-----EQSERLVKQLEPERKSTAQQL-----TRLDNLLRKEVELE- 266
Db 317 TQLLFSHEEELSKLDELEIHRINIEKLDNGLGIHYKQIQDGLQNMESQKIETMQFEK 376
Qy 267 HIAHAQAAILIAQKQKNDYNTAQSL-----RDVTAQLESVQEKYNDTAQSLRDVTAQLESEQ 321
Db 377 NLIYKQQLIIEISKLDQSLVNSKSEMTQINQLQKEI-----EILQKEKEK 428
Qy 322 EKYNDTAQSLRDVTAQLESE-QEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRD- 375
Db 429 GTLEQEQVQLQKTELLEKQMKEND---LQEKFAQLEAENSILKDEKKTLEDMLKIH 484
Qy 376 --VSAQ-----LESYKST-----LKEIEDLKLENLYQEKV-----AMAE 409
Db 485 TPVQSEERLIPLDSIKSKSDSVWEKETEILIEENEDLKQOCIQLNBEIERQNRFTSPA 544
Qy 410 KSVEDVQOQIILTAESTNOEYA--RMVQDLQNRSTLKEEIEKETSSELEKI----- 458
Db 545 KNFEVNTQEL-----QEYACLLKVKDDLEDNKNQKQLEKYSKLLKALNEELHLQRINT 598
Qy 459 -TDLKNLQROODEFRKQLEEKGRKRTAKENVMTELTMEINKRLLYBELYEKTKPFOQQ 517
Db 599 TVMKSSVFEDEKTFVAETLEMGE-VVEKDTTLMEXLEVTKREKL--ELSQRSLDLEQ 655
Qy 518 L 518
Db 656 L 656
```

RESULT 14

US-09-745-763-168

; Sequence 168, Application US/09745763

; Patent No. US20020065394A1

GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth

McCoy, John M.

LaVallie, Edward R.

Collins-Racie, Lisa A.

Evans, Cheryl

Merberg, David

Treacy, Maurice

Spaulding, Vikki

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES

ENCODING THEM

NUMBER OF SEQUENCES: 219

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genetics Institute, Inc.

STREET: 87 CambridgePark Drive

CITY: Cambridge

STATE: MA

COUNTRY: U.S.A.

ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09745763

; FILING DATE: 18-Jun-2000

; CLASSIFICATION: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Sprunger, Suzanne A.

; REGISTRATION NUMBER: 41,323

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 498-8284

; TELEFAX: (617) 876-5851

; INFORMATION FOR SEQ ID NO: 168:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 677 amino acids

; TYPE: amino acid

; STRANDEDNESS: <Unknown>

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 168:

US-09-745-763-168

Query Match 9.6%; Score 299.5; DB 10; Length 677;
Best Local Similarity 20.6%; Pred. No. 3.7e-07;
Matches 145; Conservative 153; Mismatches 224; Indels 181; Gaps 29;

Qy 24 VKQGMELKQATQKDLTESKGIV-QLEGKLVISYEKEDKCEKTELLEYIQEISCAS 82

Db 1 MKNHYVPLKV---SDMKKSHDAIIDLNRKLLDVTQYTEKKLEMEKLL-----LEN 50

Qy 83 DOVEKCKVDIAQLEBDL---KEKDRREILSKQSLSENITFSQIEDLVTKC----- 130

Db 51 DSLSK--DVSRLFTVPPEKHEKEITALKSNIVE---LKKQLSELAKKCGEDQEKIHA 104

Qy 131 -----QLLETERONLVSKDRE-----RAETLSAE-QLILTE 160

Db 105 LTSNTNLKMMNSQYVPVKTHEEVKMTLNDTLAKTNRELLDVKKKFEDINQEFVKIKDK 164

Qy 161 RLALERQ-----EYELQKQELQSLLQKQKELSLARLQQLCSF---QEEMTISE 207

Db 165 NEILKRNLENTQNIKAIEYISLAHEAKMSSLQSMRKVQDSNAEILANRYKQGEIIVT- 223

Qy 208 KNVFKELKLAELDAVO-----OKEEOS 232

Db 224 ---LHAETKAQKKELDITQECIKVKYAPIVSFEECKRFKATEKELKDQSLSEQTOKYSVS 280

Qy 233 ERLVKQLEERKSTAEQLTRLDNLLREKEVELEKIHAAHAQAAILIAQEKYNDTAQSLRDV 292

Db 281 EEEVKNKQENDKLLKEIFTQKDLRDKTVLIEK-----SHEMERALSRTDEL 329

Qy 293 TAQLESVQEKYNDTAQSLRDVTAQLESBQEKYNDTAQSLRDV-----TAQLESQEKY 345

Db 330 NKQLKDLQSKYTE---VKNVKEKLVEENAKQTSILAVQNLLQKHVPLEQVEALKKSL 385

Qy 346 NDTAQSLRDVTAQLESVQEKYNDTAQSLRDVSAQLESYKSSSTLKEIEDLKLENLYQEKV 405

Db 386 NGTIENLKE---ELKSMQRCYEKEQQTVTKLHQLLENKKNSSVPLAEHLQIKE-AFEKVV 441

Qy 406 AMAEKSVEDVQOQIILTAESTNOEYARMVDLQNL-RSTLKEEIEKEI-----TSSSLE 456

Db 442 GLIKASLREKEEEE---SQNMKEEVSKQSEVQNTQKALKKLETREVVDLSKYKATKSDLE 498

Qy 457 -KITDLKNQL---RQODEDFRKQLEEKGRKRTA---EKENVMTELTMTET-----NKWR 501

Db 499 TQISLNEKLANLRKYEEVCEVHLHAKKEISAKDENELLHFSIEQEKDKQKCRDKSL 558

Qy 502 LLYEELYEYKTPFOQLDAFAEAKQALLNEHGATQEQLNKIRDSYAQLL---GHQNLKQK 558

Db 559 TTITELQRRIQESAKQIEAKONKITELND----VERLKQALNGLSQTLYTSGNPTKROS 614

Qy 559 -----IKHVVK-----LKDENSQKSEVSKLSRSLQVKKRQKQNEL 591

Db 615 QLIDTLQHVQKSLSEQQLADADRQHOEVIAIYRTHLLSAAQGHM 657

RESULT 15

US-09-991-496-120
; Sequence 120, Application US/09991496
; Patent No. US40020169285A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skelky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Probst, Peter
; APPLICANT: Brannon, Mark
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS
; FILE REFERENCE: 210121.420C9
; CURRENT APPLICATION NUMBER: US/09/991,496
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 120
; LENGTH: 2310
; TYPE: PRT
; ORGANISM: Leishmania major and chagasi
US-09-991-496-120

Query Match 9.6%; Score 299; DB 9; Length 2310;

Best Local Similarity 20.6%; Pred. No. 1.3e-06;

Matches 149; Conservative 146; Mismatches 274; Indels 156; Gaps 26;

QY 13 NKRETKMRSMVYKQEGMELKQAT-----QKDLTFSKQIVQLECKLVISIEKEID 63
DB 557 NAEELQORLDTATQORAELEAQAARLAANAELQORLDTATQORAELEAQAARLAADRDE 616
QY 64 EKCETEKLELYIQE-ISCASDQVEKCKVDIAQLEEDLKEKREILSLKQSLSEENT-TFSK 121
DB 617 ARQOLAANAELQORLDTATQORAELEAQAARLAADGDEARQOLAANAELQORLDTATQ 576
QY 122 QIETLVKQCLLETERDNLVSKDRGAEYLS-----AEMQILTERLALERQE--- 168
DB 677 QRAELEAQAARLAADRDEARQOLAANAELQORLDTATQORAELEAQAARLAADRDEARQ 736
QY 169 -----YEKLOQK-----ELQSQ---SLLQOEK 187
DB 737 QLAANAELQORLDTATQORAELEAQAARLAADGDEARQOLAANAELQORLDTATQORA 796
QY 188 ELSARL-----QQQLCSFQEMTSEKNVF---KEELKLALAE-----DAVQOKE 229
DB 797 ELEAQAARLAADRDEARQOLAANAELQORLDTATQORAELEAQAARLAADGDEARQOLA 856
QY 230 EQSERLVQLEERKSTARQLTRNLNREKEVELEKHIAHAQAAILIAQEKYNDTAOSL 289
DB 857 ANAEELQORLDTATQORAELEAQAARLAADRDEARQOLAANAEL---QORLDTATQOR 912
QY 290 RDVTAQLESVQEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLESQEKY-NDT 348
DB 913 AELEAQAARLAADRDEARQOLAANAELQORLDTATQORAELEAQAARLAADG 965
QY 349 AQSLRDVTAQLESVQEKYNDTAQSLRDVSAQL-----ESYKSTLKEIEDL--KLENL 399
DB 966 DEARQOLAANAELQORLDTATQORAELEAQAARLAADRDEARQOLAANAELQORLDTA 1025
QY 400 T-----LQEKVAMAEKSVEDVQOQIILTA-----ESTNOEYARMVQDL 436
DB 1026 TQORAELEAQAARLAANAELQORLDTATQORAELEAQAARLAADRDEARQOLAANAEL 1085
QY 437 QN-----RSTLKEEIEKEITSSFKETIDLNQ---LRQODEDFRKOL 476
DB 1086 QORLDTATQORAELEAQAARLAANAELQORLDTATQORAELEAQAARLAANAELQORL 1145

QY 477 EEKGKRTAEKENVMTELMTWEINKWRLLYBELYEYKTKPFOOQLDAF---BAEQOALLNEHG 533
DB 1146 DTATQORAELEAQAARLAADRDEAR---QOLAANAELQORLDTATQORAELEAQAARLA 1202
QY 534 ATQEQNLKTRDSYAQLLGHQNLKQKIKHVVKLKDE-NSOLKSEVSKLSQSLYKRRKQNEIR 592
DB 1203 ANAEELQORLDTATQ--QRAELEAQAARLAADRDEARQOLAANAELQORLDTATQORAE 1260
QY 593 LOGEL 597
DB 1261 LEAQV 1265

Search completed: December 20, 2002, 15:14:50
Job time : 17.9108 secs


```
QY 61 KIDKCEKLEKLEYIQEISCASQDVQKCKVDIAQLEEDLKEKDREILSLKQSLNEINITS 120
DB 61 KIDKCEKLEKLEYIQEISCASQDVQKCKVDIAQLEEDLKEKDREILSLKQSLNEINITS 120
QY 121 KOIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALERQYEKLOQKELQSQ 180
DB 121 KOIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALERQYEKLOQKELQSQ 180
QY 181 SLLOQKELSLRLOOQLCSFQEEEMTSEKNVFEELKLALAEALDAVQOEQSERLVKOLE 240
DB 181 SLLOQKELSLRLOOQLCSFQEEEMTSEKNVFEELKLALAEALDAVQOEQSERLVKOLE 240
QY 241 EERKSTAEQLTRLDNLLREKEVELEKHIAHAQAAILIAQEKYNDTAQSLRDVTAQLESVQ 300
DB 241 EERKSTAEQLTRLDNLLREKEVELEKHIAHAQAAILIAQEKYNDTAQSLRDVTAQLESVQ 300
QY 301 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 360
DB 301 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 360
QY 361 SVQEKYNDTAQSLRDVSAQLESYKSTLKEIEDLKLLENLTLOEKVAMAEKSVEDVQOQIL 420
DB 361 SVQEKYNDTAQSLRDVTAQLESYKSTLKEIEDLKLLENLTLOEKVAMAEKSVEDVQOQIL 420
QY 421 TAESTNQEYARMVDLQNRSTLKEEIEKITSFLEKITDLKNLQROQDEDFRKOLEBK 480
DB 421 TAESTNQEYARMVDLQNRSTLKEEIEKITSFLEKITDLKNLQROQDEDFRKOLEBK 480
QY 481 KRTAEKENVMTELTMINKWRLLYBELYEKTPFOOQLDAFAEKQALLNEHGATQEQLN 540
DB 481 KRTAEKENVMTELTMINKWRLLYBELYEKTPFOOQLDAFAEKQALLNEHGATQEQLN 540
QY 541 KIRDSYAQLLGHONLKQKIKHVVKLKDENSOLKSEVSKLRSQLVKRRKQNELRLOQELDKA 600
DB 541 KIRDSYAQLLGHONLKQKIKHVVKLKDENSOLKSEVSKLRSQLVKRRKQNELRLOQELDKA 600
QY 601 LGIRHFDPSKAFCHASKENFTPLKEGPNCC 631
DB 601 LGIRHFDPSKAFCHASKENFTPLKEGPNCC 631

RESULT 2
US-08-477-831C-2
; Sequence 2, Application US/08477831C
; Patent No. 6429291
; GENERAL INFORMATION:
; APPLICANT: TURLEY, EVA A.
; APPLICANT: SHUWEN, ZHANG
; APPLICANT: ENTWISTLE, JOYCELYN
; TITLE OF INVENTION: HVALURONAN RECEPTOR PROTEIN
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10020-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Rel. #1.0, ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477.831C
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: PIERRI, MARGARET A.
; REGISTRATION NUMBER: 30,709
; REFERENCE/DOCKET NUMBER: SIM-10
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 606 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: /desc = "RHAMM I protein"
US-08-477-831C-2
```

```
Query Match 95.1%; Score 2952.5; DB 4; Length 606;
Best Local Similarity 95.7%; Pred. No. 3.7e-171;
Matches 604; Conservative 2; Mismatches 0; Indels 25; Gaps 1;
```

```
QY 1 MRALSLEMLKLRNKRRETKMRSMVMVQEGMELKLQATQKDLTESKGIYVQLEGLYSIEKE 60
DB 1 MRALSLEMLKLRNKRRETKMRSMVMVQEGMELKLQATQKDLTESKGIYVQLEGLK----- 54
QY 61 KIDKCEKLEKLEYIQEISCASQDVQKCKVDIAQLEEDLKEKDREILSLKQSLNEINITS 120
DB 55 -----CASDQVEKCKVDIAQLEEDLKEKDREILSLKQSLNEINITS 95
QY 121 KOIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALERQYEKLOQKELQSQ 180
DB 96 KOIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALERQYEKLOQKELQSQ 155
QY 181 SLLOQKELSLRLOOQLCSFQEEEMTSEKNVFEELKLALAEALDAVQOEQSERLVKOLE 240
DB 156 SLLOQKELSLRLOOQLCSFQEEEMTSEKNVFEELKLALAEALDAVQOEQSERLVKOLE 215
QY 241 EERKSTAEQLTRLDNLLREKEVELEKHIAHAQAAILIAQEKYNDTAQSLRDVTAQLESVQ 300
DB 216 EERKSTAEQLTRLDNLLREKEVELEKHIAHAQAAILIAQEKYNDTAQSLRDVTAQLESVQ 275
QY 301 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 360
DB 276 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 335
QY 361 SVQEKYNDTAQSLRDVSAQLESYKSTLKEIEDLKLLENLTLOEKVAMAEKSVEDVQOQIL 420
DB 336 SVQEKYNDTAQSLRDVTAQLESYKSTLKEIEDLKLLENLTLOEKVAMAEKSVEDVQOQIL 395
QY 421 TAESTNQEYARMVDLQNRSTLKEEIEKITSFLEKITDLKNLQROQDEDFRKOLEBK 480
DB 396 TAESTNQEYARMVDLQNRSTLKEEIEKITSFLEKITDLKNLQROQDEDFRKOLEBK 455
QY 481 KRTAEKENVMTELTMINKWRLLYBELYEKTPFOOQLDAFAEKQALLNEHGATQEQLN 540
DB 456 KRTAEKENVMTELTMINKWRLLYBELYEKTPFOOQLDAFAEKQALLNEHGATQEQLN 515
QY 541 KIRDSYAQLLGHONLKQKIKHVVKLKDENSOLKSEVSKLRSQLVKRRKQNELRLOQELDKA 600
DB 516 KIRDSYAQLLGHONLKQKIKHVVKLKDENSOLKSEVSKLRSQLVKRRKQNELRLOQELDKA 575
QY 601 LGIRHFDPSKAFCHASKENFTPLKEGPNCC 631
DB 576 LGIRHFDPSKAFCHASKENFTPLKEGPNCC 606
```

```
RESULT 3
US-08-402-217A-3
; Sequence 3, Application US/08402217A
; Patent No. 5587301
; GENERAL INFORMATION:
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Wilde, Craig G.
; APPLICANT: Seilhamer, Jeffrey J.
; TITLE OF INVENTION: HVALURONAN RECEPTOR EXPRESSED IN HUMAN
; TITLE OF INVENTION: UMBILICAL VEIN ENDOTHELIAL CELLS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
```



```
Db 301 LEKITDLKNQLRQODEDFRKQLEEKGRKTAENKENVMTLMEINKWRLLYEELYEKTSPF 360
QY 515 QOOLDAFAEAKQALLNEHGATQEQNLKIRDSYAQLLGHQNLKOKIKHVVYKLDKENSOLKS 574
Db 361 QOOLDAFAEAKQALLNEHGATQEQNLKIRDSYAQLLGHQNLKOKIKHVVYKLDKENSOLKS 420
QY 575 EVSKLRSQLVKRRKQNELRQGLDQALGIRHFDPSKAFCHASKENFTPLKEGNPNCC 631
Db 421 EVSKLRSQLVKRRKQNELRQGLDQALGIRHFDPSKAFCHASKENFTPLKEGNPNCC 477

RESULT 5
US-08-995-654-3
; Sequence 3, Application US/08995654
; Patent No. 6025138
; GENERAL INFORMATION:
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Wilde, Craig
; APPLICANT: Seilhamer, Jeffrey
; TITLE OF INVENTION: HVALURONAN RECEPTOR EXPRESSED IN HUMAN
; TITLE OF INVENTION: UMBILICAL VEIN ENDOTHELIAL CELLS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/995,654
; FILING DATE: December 22, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/700,178
; FILING DATE: August 20, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/402,217
; FILING DATE: March 10, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0028-2 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 477 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-995-654-3

Query Match 75.8%; Score 2354; DB 3; Length 477;
Best Local Similarity 100.0%; Pred. No. 4e-135;
Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 MOILTERLALEREQYEKLOQKELQSQLSLQOQKELSAKELQSQLSQFQEMTSEKNVFKEE 214
Db 1 MOILTERLALEREQYEKLOQKELQSQLSLQOQKELSAKELQSQLSQFQEMTSEKNVFKEE 60
QY 215 LKLAELDAVQOKEEQSERLVKQLEERKSTAEQLTDLNLLREKEVELEKHHIAHAQA 274
Db 61 LKLAELDAVQOKEEQSERLVKQLEERKSTAEQLTDLNLLREKEVELEKHHIAHAQA 120
```

```
QY 275 ILIAQEKYNDTAQSLRDVTAQLESVOEKYNDTAQSLRDVTAQLESQEQEKYNDTAQSLRDV 334
Db 121 ILIAQEKYNDTAQSLRDVTAQLESVOEKYNDTAQSLRDVTAQLESQEQEKYNDTAQSLRDV 180
QY 335 TAOLESQEKYNDTAQSLRDVTAQLESVOEKYNDTAQSLRDVTAQLESYKSSSTLKEIEDL 394
Db 181 TAOLESQEKYNDTAQSLRDVTAQLESVOEKYNDTAQSLRDVTAQLESYKSSSTLKEIEDL 240
QY 395 KLENLTLOEKVAMAEKSVEDVQOQILTAESTNQEYARMVQDLQNRSTLKEEIEIKETTSF 454
Db 241 KLENLTLOEKVAMAEKSVEDVQOQILTAESTNQEYARMVQDLQNRSTLKEEIEIKETTSF 300
QY 455 LEKITDLKNQLRQODEDFRKQLEEKGRKTAENKENVMTLMEINKWRLLYEELYEKTSPF 514
Db 301 LEKITDLKNQLRQODEDFRKQLEEKGRKTAENKENVMTLMEINKWRLLYEELYEKTSPF 360
QY 515 QOOLDAFAEAKQALLNEHGATQEQNLKIRDSYAQLLGHQNLKOKIKHVVYKLDKENSOLKS 574
Db 361 QOOLDAFAEAKQALLNEHGATQEQNLKIRDSYAQLLGHQNLKOKIKHVVYKLDKENSOLKS 420
QY 575 EVSKLRSQLVKRRKQNELRQGLDQALGIRHFDPSKAFCHASKENFTPLKEGNPNCC 631
Db 421 EVSKLRSQLVKRRKQNELRQGLDQALGIRHFDPSKAFCHASKENFTPLKEGNPNCC 477

RESULT 6
US-08-402-217A-2
; Sequence 2, Application US/08402217A
; Patent No. 5587301
; GENERAL INFORMATION:
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Wilde, Craig G.
; APPLICANT: Seilhamer, Jeffrey J.
; TITLE OF INVENTION: HVALURONAN RECEPTOR EXPRESSED IN HUMAN
; TITLE OF INVENTION: UMBILICAL VEIN ENDOTHELIAL CELLS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/402,217A
; FILING DATE: 10-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Luther, Barbara J.
; REGISTRATION NUMBER: 33954
; REFERENCE/DOCKET NUMBER: PF-002808
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-852-0195
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 351 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-402-217A-2

Query Match 38.8%; Score 1205; DB 1; Length 351;
Best Local Similarity 59.6%; Pred. No. 6.9e-66;
Matches 258; Conservative 29; Mismatches 62; Indels 84; Gaps 1;

QY 155 MOILTERLALEREQYEKLOQKELQSQLSLQOQKELSAKELQSQLSQFQEMTSEKNVFKEE 214
||| :: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 1 MNLKQKFILEQOEREKLOQKELQDSLQOQKELSSSLHQKLCFSQOEMAKEKNLFEE 60
QY 215 LKLAELDAVQOKEQERLQKLEERKSTAEQOLTRLDNLLREKEVELEKHAHAQA 274
Db 61 LKOTDELKLOQKEQERLQKLEERKSTAEQOLTRLDNLLREKEVELEKHAHAQA 120
QY 275 ILIAQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDV 334
Db 121 TULLBEKYDSMVQSLDVT----- 139
QY 335 TAQLESEQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDV 394
Db 140 -----AQFESYKALTASIEDL 156
QY 395 KLENLTLOEKVAMAEKSVEDVQOQILTAESTNOEYARMVQDQNLQKSTLKEEIKTTSF 454
Db 157 KLENSLQEKVAKAGNAEDVQHQILATESSNOEYARMVQDQNLQKSTLKEEIKTTSF 216
QY 455 LEKITDLNQLRQODEDFRKQLEEKGRKTAENYMTLTMEINKWRLLYEELYEYKTKPF 514
Db 217 LOKITDLNQLRQODEDFRKQLEEKGRKTAENYMTLTMEINKWRLLYEELYEYKTKPF 276
QY 515 QOOLDAFAEAKOALLNEHCATQEOQNLKIRDSYAQLLGHONLKQKIKHVYKLDKENSOLKS 574
Db 277 QLOLDAFEVEKQALLNEHGAQEQNLKIRDSYAQLLGHONLKQKIKHVYKLDKENSOLKS 336
QY 575 EVSKLRSOLVKRK 587
Db 337 EVSKLRCQLAKKK 349

RESULT 7

US-08-700-178-2
; Sequence 2, Application US/08700178
; Patent No. 5783669 5700912
; Patent No. 5783669 5700912
; GENERAL INFORMATION:
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Wilde, Craig G.
; APPLICANT: Seilhamer, Jeffrey J.
; TITLE OF INVENTION: HYALURONAN RECEPTOR EXPRESSED IN HUMAN
; TITLE OF INVENTION: UMBILICAL VEIN ENDOTHELIAL CELLS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/700,178
; FILING DATE: August 20, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/402,217
; FILING DATE: March 10, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0028-1 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 351 amino acids
; TYPE: amino acid

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-700-178-2
Query Match 38.8%; Score 1205; DB 1; Length 351;
Best Local Similarity 59.6%; Pred. No. 6.9e-66;
Matches 258; Conservative 29; Mismatches 62; Indels 84; Gaps 1;
QY 155 MOILTERLALEREQEKLOQKELQDSLQOQKELSSSLHQKLCFSQOEMAKEKNLFEE 214
Db 1 MNLKQKFILEQOEREKLOQKELQDSLQOQKELSSSLHQKLCFSQOEMAKEKNLFEE 60
QY 215 LKLAELDAVQOKEQERLQKLEERKSTAEQOLTRLDNLLREKEVELEKHAHAQA 274
Db 61 LKOTDELKLOQKEQERLQKLEERKSTAEQOLTRLDNLLREKEVELEKHAHAQA 120
QY 275 ILIAQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDV 334
Db 121 TULLBEKYDSMVQSLDVT----- 139
QY 335 TAQLESEQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDV 394
Db 140 -----AQFESYKALTASIEDL 156
QY 395 KLENLTLOEKVAMAEKSVEDVQOQILTAESTNOEYARMVQDQNLQKSTLKEEIKTTSF 454
Db 157 KLENSLQEKVAKAGNAEDVQHQILATESSNOEYARMVQDQNLQKSTLKEEIKTTSF 216
QY 455 LEKITDLNQLRQODEDFRKQLEEKGRKTAENYMTLTMEINKWRLLYEELYEYKTKPF 514
Db 217 LOKITDLNQLRQODEDFRKQLEEKGRKTAENYMTLTMEINKWRLLYEELYEYKTKPF 276
QY 515 QOOLDAFAEAKOALLNEHCATQEOQNLKIRDSYAQLLGHONLKQKIKHVYKLDKENSOLKS 574
Db 277 QLOLDAFEVEKQALLNEHGAQEQNLKIRDSYAQLLGHONLKQKIKHVYKLDKENSOLKS 336
QY 575 EVSKLRSOLVKRK 587
Db 337 EVSKLRCQLAKKK 349
RESULT 8
US-08-995-654-2
; Sequence 2, Application US/08995654
; Patent No. 6025138
; GENERAL INFORMATION:
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Wilde, Craig
; APPLICANT: Seilhamer, Jeffrey
; TITLE OF INVENTION: HYALURONAN RECEPTOR EXPRESSED IN HUMAN
; TITLE OF INVENTION: UMBILICAL VEIN ENDOTHELIAL CELLS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/995,654
; FILING DATE: December 22, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/700,178
; FILING DATE: August 20, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/402,217

[illegible]


```
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6379912el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1017
; CURRENT APPLICATION NUMBER: US/09/723,262
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: US 09/572,191
; PRIOR FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1388
; TYPE: PRT
; ORGANISM: Human
US-09-723-262-2

Query Match      12.1%; Score 375; DB 4; Length 1388;
Best Local Similarity 22.8%; Pred. No. 4.2e-15;
Matches 165; Conservative 142; Mismatches 230; Indels 188; Gaps 28;

Qy 3 ALSLEMLKLRNKRTRKMRSMVMVKGEMELKQATQKDLTESGKIVQLEGKLVSTIEKEKI 62
Db 714 AISEELRTVQ-EQMSALQAKLDEEHEKNLKLQOHVVDKLEHHSTQMOEL-----FSSERI 766

Qy 63 DEKCETEKLL-----EYIQEISCASD-----QVEKCKVDIAQLPEDLKEDREILSLKQ 111
Db 767 DWTQOQELLNOLNLYLEKQLOQTQTNFLKSEVHDLRVLHSDAKELSSVKLEYSSEFT 826

Qy 112 SLEENITFSKQIE-DLTVKQCL-----LETER-----DNL----- 140
Db 827 NOEKE--FNKLSERHMHVQLQDLNLRLENKLESKACLODSYDNLQETIMKEFIDQLSRN 884

Qy 141 VSKDRERAEITSAEMQIILTERIALERQYEKL-----QKELQSQSL-----QOKE 188
Db 885 LQNFKENETLKSDLNLMLELAEKERNKLSLOFEDKENSKEILKVLAEVQRQKQ 944

Qy 189 LSARLQOQLCSFOEBMTSEKNVFKELALAEALDAVQKEQSER-----LVKOLEER 243
Db 945 ETAKCEQOMAKVQK-----LEESLLATEKVISSLEKSRSDSKVVDLMNQIQLR 995

Qy 244 KSTABQLTRLDNLLRE-KEVELEKHIA---AHAQAAILIAOEKYN--DTAQS----- 289
Db 996 SSVCKETETIDTLKQELKDINCKYNSALVDREESRVLIKQEVVDLDLKETLRILSED 1055

Qy 290 --RDVTAQ--LESVQEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLESEQEKYN 346
Db 1056 TERDMLCEDLAHATEQLNMLTEASKKHSGLLSQAQOEELTKREALIQELQHLNOKKEEVE 1115

Qy 347 DTAQSLRDVTAQLESVQEKYNDTAQS-----LRDVSQALESYKSS----- 386
Db 1116 QKKNEYNFKMRQLEHVMSAAEDPQSPKTPPHFQTHLAKLLETQEQEIEDGRASKTSLH 1175

Qy 387 -----TLKEIEDLKLENLTLOEKVAMAESVEDVQOQILTAEST 425
Db 1176 LVTKLINEDREVKNABILRMKEQLREMNLRLESQOLIEKNWLLQGLDIDIKRQ---KENS 1232

Qy 426 NOEYARMVQDQLNRLTKLEEIKE-----ITSSFLEKITDLKNO-----LRQO 468
Db 1233 DQNH----PDNQLKNEQESIKERLAKSKIIVEEMLKMKADLEEVQSALYNKEMCLRMT 1288

Qy 469 DEDFRKQLEEGKRTAEKENVMTLTMTINKWRLLYELYEKTKPFQOQLDAFAEAKAL 528
Db 1337 AEENG-----KLVGHNQLHQTQYVYVLKKNVNRVLAETTEKLAENVFLKE 1382

Qy 586 RQNE 590
:|:|:
```

```
Db 1383 KKRSE 1387

RESULT 15
US-09-723-219-2
; Sequence 2, Application US/09723219
; Patent No. 6391613
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6391613el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1017
; CURRENT APPLICATION NUMBER: US/09/723,219
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: US 09/572,191
; PRIOR FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1388
; TYPE: PRT
; ORGANISM: Human
US-09-723-219-2

Query Match      12.1%; Score 375; DB 4; Length 1388;
Best Local Similarity 22.8%; Pred. No. 4.2e-15;
Matches 165; Conservative 142; Mismatches 230; Indels 188; Gaps 28;

Qy 3 ALSLEMLKLRNKRTRKMRSMVMVKGEMELKQATQKDLTESGKIVQLEGKLVSTIEKEKI 62
Db 714 AISEELRTVQ-EQMSALQAKLDEEHEKNLKLQOHVVDKLEHHSTQMOEL-----FSSERI 766

Qy 63 DEKCETEKLL-----EYIQEISCASD-----QVEKCKVDIAQLPEDLKEDREILSLKQ 111
Db 767 DWTQOQELLNOLNLYLEKQLOQTQTNFLKSEVHDLRVLHSDAKELSSVKLEYSSEFT 826

Qy 112 SLEENITFSKQIE-DLTVKQCL-----LETER-----DNL----- 140
Db 827 NOEKE--FNKLSERHMHVQLQDLNLRLENKLESKACLODSYDNLQETIMKEFIDQLSRN 884

Qy 141 VSKDRERAEITSAEMQIILTERIALERQYEKL-----QKELQSQSL-----QOKE 188
Db 885 LQNFKENETLKSDLNLMLELAEKERNKLSLOFEDKENSKEILKVLAEVQRQKQ 944

Qy 189 LSARLQOQLCSFOEBMTSEKNVFKELALAEALDAVQKEQSER-----LVKOLEER 243
Db 945 ETAKCEQOMAKVQK-----LEESLLATEKVISSLEKSRSDSKVVDLMNQIQLR 995

Qy 244 KSTABQLTRLDNLLRE-KEVELEKHIA---AHAQAAILIAOEKYN--DTAQS----- 289
Db 996 SSVCKETETIDTLKQELKDINCKYNSALVDREESRVLIKQEVVDLDLKETLRILSED 1055

Qy 290 --RDVTAQ--LESVQEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLESEQEKYN 346
Db 1056 TERDMLCEDLAHATEQLNMLTEASKKHSGLLSQAQOEELTKREALIQELQHLNOKKEEVE 1115

Qy 347 DTAQSLRDVTAQLESVQEKYNDTAQS-----LRDVSQALESYKSS----- 386
Db 1116 QKKNEYNFKMRQLEHVMSAAEDPQSPKTPPHFQTHLAKLLETQEQEIEDGRASKTSLH 1175

Qy 387 -----TLKEIEDLKLENLTLOEKVAMAESVEDVQOQILTAEST 425
Db 1176 LVTKLINEDREVKNABILRMKEQLREMNLRLESQOLIEKNWLLQGLDIDIKRQ---KENS 1232

Qy 426 NOEYARMVQDQLNRLTKLEEIKE-----ITSSFLEKITDLKNO-----LRQO 468
Db 1233 DQNH----PDNQLKNEQESIKERLAKSKIIVEEMLKMKADLEEVQSALYNKEMCLRMT 1288

Qy 469 DEDFRKQLEEGKRTAEKENVMTLTMTINKWRLLYELYEKTKPFQOQLDAFAEAKAL 528
Db 1289 DEVERTOTLE-SKAFQEQELRSKL-----EENYERERTSOEMELRQVQVCL 1336
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:09:43 ; Search time 31.6431 Seconds
(without alignments)
4108.822 Million cell updates/sec

Title: US-09-685-010-48
Perfect score: 3104
Sequence: 1 MRALSLEMLKLNKRKRETKMR.....FCHASKENFTPLKEGNPNCC 631

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phase.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3048	98.2	836	11 Q8VDR2	Q8vdr2 mus musculu
2	2374.5	76.5	713	11 Q9WUF7	Q9wuf7 rattus norv
3	2338.5	75.3	476	11 Q920Z6	Q920z6 mus sp. hya
4	616.5	19.9	240	6 Q8SPR5	Q8spr5 bos taurus
5	437.5	14.1	1598	11 Q922D2	Q922d2 mus musculu
6	431.5	13.9	1790	3 Q07380	Q07380 saccharomyc
7	428.5	13.8	1937	6 Q9TV62	Q9tv62 sus scrofa
8	425	13.7	1940	6 Q9BE41	Q9be41 bos taurus
9	423	13.6	1939	6 Q9TV63	Q9tv63 sus scrofa
10	418	13.5	1944	13 Q9DGM5	Q9dgm5 gallus gall
11	417	13.4	1938	6 Q9BE40	Q9be40 bos taurus
12	415	13.4	1410	4 Q14221	Q14221 homo sapien
13	414	13.3	1411	4 Q15075	Q15075 homo sapien
14	413	13.3	1939	13 Q9PTY2	Q9pty2 gallus gall
15	411	13.2	1939	6 Q9TV61	Q9tv61 sus scrofa
16	408.5	13.2	1690	5 O44929	O44929 drosophila

17	405	13.0	1941	13 Q9DCM4	Q9dcm4 gallus gall
18	402.5	13.0	1935	13 Q90ZE5	Q90ze5 brachydanio
19	402	13.0	1938	13 Q9IBD7	Q9ibd7 seriola dum
20	401.5	12.9	1933	13 Q90337	Q90337 cyprinus ca
21	399.5	12.9	1950	5 Q26080	Q26080 placopecten
22	398.5	12.8	1941	5 Q26079	Q26079 placopecten
23	397	12.8	1368	13 Q9PTD7	Q9pty7 xenopus lae
24	396.5	12.8	1690	5 Q9VJE5	Q9vje5 drosophila
25	395.5	12.7	1937	13 Q9IBD4	Q9ibd4 gallus gall
26	395	12.7	1932	13 Q98TQ4	Q98tq4 notochenia
27	395	12.7	1938	6 Q9GJP9	Q9gjp9 oryctolagus
28	393	12.7	1463	5 Q9GYZ0	Q9gyz0 strongyloce
29	391	12.6	2954	13 Q42263	Q42263 xenopus lae
30	390.5	12.6	1092	13 Q90338	Q90338 cyprinus ca
31	390.5	12.6	1931	13 Q42352	Q42352 cyprinus ca
32	390	12.6	1738	5 Q76329	Q76329 dictyosteli
33	389	12.5	1929	13 Q98TQ6	Q98tq6 notochenia
34	388.5	12.5	1935	4 Q9HID5	Q9hid5 homo sapien
35	388	12.5	1958	5 Q96062	Q96062 dugesia jap
36	387.5	12.5	1583	4 Q15045	Q15045 homo sapien
37	387	12.5	1960	11 Q8VDD5	Q8vdd5 mus musculu
38	386.5	12.5	1935	6 Q9BE39	Q9be39 bos taurus
39	386	12.4	1930	13 Q9DGD5	Q9dgd5 pennahia ar
40	384.5	12.4	2139	5 Q07569	Q07569 entamoeba h
41	384	12.4	1936	13 Q90YF6	Q90yfe paracirrhith
42	384	12.4	1979	5 Q96133	Q96133 plasmodium
43	383.5	12.4	1931	13 Q91OC5	Q91oc5 gallus gall
44	383.5	12.4	1935	6 Q9GKRL	Q9gkrl sus scrofa
45	381	12.3	1935	11 Q91283	Q91283 mus musculu

ALIGNMENTS

RESULT 1

Q8VDR2 ID Q8VDR2 PRELIMINARY; PRT; 836 AA.
AC Q8VDR2;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Similar to hyaluronan mediated motility receptor (RHAMM).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC021427; AAH21427.1; -
KW Receptor.
SQ SEQUENCE 836 AA; 96670 MW; B9BDDA22EBE6A652 CRC64;

Query Match	98.2%	Score 3048;	DB 11;	Length 836;
Best Local Similarity	93.0%	Pred. No. 2.8e-101;		
Matches 626;	Conservative 2;	Mismatches 3;	Indels 42;	Gaps 1;
Qy 1	MRALSLEMLKLNKRKRETKMRSMVQEGMELKQATQKDLTSGKIVOLEGKLVSIKIE 60			
Db 164	MRALSLEMLKLNKRKRETKMRSMVQEGMELKQATQKDLTSGKIVOLEGKLVSIKIE 223			
Qy 61	KIDKCETEKLELEYTOEISCSASDQVECKVDIAQLEEDLKEKDREILSLKQSLSENIPTS 120			
Db 224	KIDKCETEKLELEYTOEISCSASDQVECKVDIAQLEEDLKEKDREILSLKQSLSENIPTS 283			
Qy 121	KQIEDLTVCQLLETERNLVSKDRERATLSAEQMILTERLALERQVEYKIQKELQKELQSO 180			
Db 284	KQIEDLTVCQLLETERNLVSKDRERATLSAEQMILTERLALERQVEYKIQKELQKELQSO 343			
Qy 181	SLLQKEKELSLARLQQLCSFQEMTSEKNVFKKELKALAEALDAVQKKEQSERLVKQLE 240			
Db 344	SLLQKEKELSLARLQQLCSFQEMTSEKNVFKKELKALAEALDAVQKKEQSERLVKQLE 403			

Query Match	75.3%	Score 2338.5	DB 11	Length 476
Best Local Similarity	99.8%	Pred. No. 2.5e-76		
Matches 476	Conservative 0	Mismatches 0	Indels 1	Gaps 1
QY 155	MQILTERLALERQYKQKQELQSQSLQOQKELSLARLQOQLCSFOEEMTSEKNVPKEE	214		
DB 1	MQILTERLALERQYKQKQELQSQSLQOQKELSLARLQOQLCSFOEEMTSEKNVPKEE	60		
QY 215	LKLALAEALDAVQOKEQSERLVKQLEERKSTAEQLTRLDNLLREKEVELEKHHIAHAQA	274		
DB 61	LKLALAEALDAVQOKEQSERLVKQLEERKSTAEQLTRLDNLLREKEVELEKHHIAHAQA	120		
QY 275	ILTAQEKYNDTAQSLRDVDTAQLESVQVKYNDTAQSLRDVDTAQLESPOEKYNDTAQSLRDV	334		
DB 121	ILTAQEKYNDTAQSLRDVDTAQLESVQVKYNDTAQSLRDVDTAQLESPOEKYNDTAQSLRDV	180		
QY 335	TAQLESPOEKYNDTAQSLRDVDTAQLESVQVKYNDTAQSLRDVDTAQLESVQVKYNDTAQSLRDV	394		
DB 181	TAQLESPOEKYNDTAQSLRDVDTAQLESVQVKYNDTAQSLRDVDTAQLESVQVKYNDTAQSLRDV	240		
QY 395	KLENLTLOEKVMAEKSVEDVQOQILTAESTNQEIYARMVQDLQNRSLTKEEIKETSSSF	454		
DB 241	KLENLTLOEKVMAEKSVEDVQOQILTAESTNQEIYARMVQDLQNRSLTKEEIKETSSSF	300		
QY 455	LEKITDLKQNLROQDEDFRQLEBKGRKRTAKENVMVTELTMEINKNRLLYEELYEYKTKPF	514		
DB 301	LEKITDLKQNLROQDEDFRQLEBKGRKRTAKENVMVTELTMEINKNRLLYEELYEYKTKPF	360		
QY 515	QQQLDAFEAKQALLNEHGATQEQNLKIRSYAQLLGHQNLKQIKHVVKLKDENSOLKS	574		
DB 361	-QQQLDAFEAKQALLNEHGATQEQNLKIRSYAQLLGHQNLKQIKHVVKLKDENSOLKS	419		
QY 575	EVSKLSQLVKRRKQNELRLOGLDQKALGIRHFDPSKAFCHASKENFTPLKEGPNPCC	631		
DB 420	EVSKLSQLVKRRKQNELRLOGLDQKALGIRHFDPSKAFCHASKENFTPLKEGPNPCC	476		
RESULT 4				
Q8SPR5	PRELIMINARY;	PRT;	240 AA.	
ID	Q8SPR5	PRELIMINARY;	PRT;	240 AA.
AC	Q8SPR5			
DT	01-JUN-2002 (TEMBLrel. 21, Created)			
DT	01-JUN-2002 (TEMBLrel. 21, Last sequence update)			
DT	01-JUN-2002 (TEMBLrel. 21, Last annotation update)			
DE	Hyaluronic acid-mediated motility receptor (Fragment).			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI_TaxID=9913;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Krebs O., Alt H., Prella K., Stojkovic M.			
RT	"Receptor for hyaluronic acid mediated motility in cow (Bos taurus)			
RL	ovary cortex."			
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF499008; AAM1828.1; "			
KW	Receptor.			
FT	NON_TER	1		
FT	NON_TER	240		
FT	NON_TER	240		
SQ	SEQUENCE	240 AA;	E0BFED9F3E772E32	CRC64;

```

Query Match      19.9%; Score 616.5; DB 6; Length 240;
Best Local Similarity 74.4%; Pred. No. 1.8e-15;
Matches 128; Conservative 24; Mismatches 19; Indels 1; Gaps 1;

QY 1 MRAISLEMLKRNKRETKMRSMVMVQSGMELKLTQATQKDLTESGKGVQLVLEGKLVLSIEKE 60
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 69 LRILSLEMLKRNKRETKMRSMMAQSGMEVKLQLTQKNLVSESQEKIVQLVLEGKLVLSIEKE 128

QY 61 KIDEKCTEKLLEYIQBIESCASDOVEKCKVDIAQLEDLKEKREIILSLKOSLENT-F 119
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 129 KIDEKSTETKLLEYIEISCASDOVEKVKLDIAQLENLKEKNHEVLSLKOSLENTIL 188

```

```

Qy 120 SKQIEDTVKCOLLETERDNLVSKDRRAEFTLSAEMQILTERLALEROEYEK 171
      |||||: 1 1 1 1 1:|:|:| 1 1:|:| 1 1 1 1:|:|:|
Db 189 SKQEDMKSQYVLEEKEDLINRNRDERDNLSEIQSLKERSILEKQEHHE 240

RESULT 5
Q922D2 PRELIMINARY; PRT; 1598 AA.
AC Q922D2:
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to myosin, heavy polypeptide 2, skeletal muscle,
DE adult.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
[1]
RN SEQUENCE FROM N.A.
RP
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC008538; AA08538.1; -.
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR001609; myosin_head.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR PROSITE; PS50096; IQ; 1.
SQ SEQUENCE 1598 AA; 183083 MW; DE896B343464A479 CRC64;

```

```

Query Match      14.1%   Score 437.5; DB 11; Length 1598;
Best Local Similarity 24.0%; Pred. No. 2.le-08;
Matches 180; Conservative 134; Mismatches 249; Indels 187; Gaps 29;
```

Qy	4	LSELMKLRNREFTMRSMWYQEGMEKLQAOTOKDLTESKGKIVQLGCKLVSTIEKEKID	63
Db	839	LFFIKIPLLKSATTEKENATMKEE-----FOKTDDDLAKSAKRKELEPKVMVSLLKEND	893
Qy	64	EKCETEKLEVIIQETSCASDOVEKCV-----DIAOLEEDLKKDEIRILLSKQSLEEN	116
Db	894	LQLQVAEGAELADAEPCDGLINTKIQLAKIIEVTERAEEDEINELTAKRKKLEDE	953
Qy	117	IT-FSKQIEDITVRQCQLLETERRDLNVSKDRERAETLSAEMQILFRLALEREQEYKLOQK	175
Db	954	CSELKKDIDDELTLAKVKEKEKHATENK----VNMLTEEMAGLDETIAKLTKEKKAQEA	1009
Qy	176	ELQSGSLQQOBE-----LSARIQQOLCSFGEMTSEKNV-----FKUELKA	218
Db	1010	HQOTIDDDLQAEBEDKVNTLTAKIKILEGOVDLEGSEGLEKRLMDLERAKRKLGGDKLA	1069
Qy	219	LAE-LDAVQOEESQSERLVK-----QLEEE-----RKSTAQLFRDLNLLREK	260
Db	1070	QESIMDIENEKOQDLDERLUKKFEPMNSIQSIEDEQATGIQLKKIKELQARIEL--EE	1127
Qy	261	EVELEKHTAAHAQAILLIAQEKYNDTAQSLRDVTQAQLESVQEKYNDTAQSLRDVTQAQLESE	320
Db	1128	EIEAER--ASRAKA-----EKQRS-----DLRSRELEISERLEANGA---TSAQITEMN	1171
Qy	321	QEKYNDTAQSLRDVTQAQLESQEKYNDTAQSLRDVTQAQLESVQEKYNDTAQSLRDVSQAQL	380
Db	1172	KKREAEPQKMRRD---LEEATLQHEATAATLRKKHA--DSVAE-LGEQIDNLQRVKQL	1224
Qy	391	ESYSKSTLIKETIEDLKLENLTLOEKVAMAEKSVEDVQQQILTAESTNOEIYARMWOOL--ONR	439
Db	1225	EKESKMEWIDDLASNVTETYSKAGNIUECMKRTLEDQVSKLSSKEEQEQQRIINDLTSQR	1284


```
Db 892 QVQAEADGLADAEERCDOLIKTQLEAKIKVENTERADEEBEINAEITAKRKKLEDECS 951
Qy 119 FSKQIEDLTVKCOLLETERDNVSKDRERAETLSAEMOILTERLALERQVEYKLOQKELQ 178
Db 952 LKKDIDDLLELTLAKVEKKEHATENK-----VKNLTEEMAGLDENIAKTKEKALQEAHQ 1007
Qy 179 SQSLQOEKE-----LSARLOOQLCSFOEEMTSEKNV-----FRELKAL-A 220
Db 1008 TLDDLOAEEDVNTLTAKTKTLEQVDDLEGSLEQEKLRMDLERAKRKKLEGLKLAQES 1067
Qy 221 ELDAVOQKEEBSERLVKQ-----LE-----EERKSTAEQOLTRLDNLLREKEVELEKHIAAH 271
Db 1068 TMDIENDKQQLDEKLKKEFEFMSNLQSKIEDEQALAMOLKQIKELQARTAELEEEIEAE 1127
Qy 272 AQAIILIAOEKYNDTAQSLRDVTAQLESVOEKYNDTAQSLRDVTAQLESQEKYNDTAQSL 331
Db 1128 RASRAKAEKQSDLSRELEIEISERLE-----EAGGATSQIEMNKKREAEFQKMR 1177
Qy 332 RDVTAQLESQEKYNDTAQSLRDVTAQLESVOEKYNDTAQSLRDVSAQLESYKSSTLKEI 391
Db 1178 RD-----LEEATLQHEATAAALRKKHA--DSVAE-LGEQIDNLRVKQKLEKEKSELKWEI 1230
Qy 392 EDLKLNTLTQEKVMAEKSVEDVQOQILTAESTNQEYARVMVQDLQNRSTLKEEBIKIT 451
Db 1231 DDLASNMETVSKAGNLEKMCRTLEDQLSEYKTRKEEHQRLINELSAQKARLQTESGEFS 1290
Qy 452 -----SSFEKITDLKNLROQ-----DED-FRKOLE 477
Db 1291 QRLDEKEALVSOLSRGKQAFQOIEELARQLEETKAKSALAHAVQSSRHDCDLLREQYE 1350
Qy 478 EKGKRTAEKENVMTLTMEINKRLLY-----EELYEKTQPFQOQLD----- 519
Db 1351 EEQEAKEALQRAKSEANSEVAQWRTKYETDAIQRTEELEEAQKLAQRLQDAEERHVAVN 1410
Qy 520 -----AFEAKEQALLNE-----HGATQEOQLN 540
Db 1411 AKCASLEKTKQRLQNEVEDLMDVRSNAACAALDKKQNFDKILAENKHKRYEETQAELE 1470
Qy 541 -----KIRDSYAQLLG-----HONLKOKI-----KHVVKLKD 567
Db 1471 ASQKESRSLSTELFKVKNAYEESLDQLETLRKKNLQOEISDLTEQTAEGGKHIELEK 1530
Qy 568 ENSQLKSEVSKLRSQVRRKQNELRLOGELDKALGIR 604
Db 1531 VKKQIEQEKSELQAL-----EEAEASLEHEEGKILRIQ 1564

RESULT 8
Q9BE41 ID Q9BE41 PRELIMINARY; PRT; 1940 AA.
AC Q9BE41;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Myosin heavy chain 2a.
GN MYHC-2A.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HOLSTEIN; TISSUE-SKELETAL MUSCLE;
RA Chikuni K., Muroya S., Nakajima I.;
RT "Sequencing of the bovine myosin heavy chain isoforms."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB059398; BAB40920.1; -.
DR HSSP; P13538; 2MYS
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR001609; myosin_head.
DR InterPro; IPR004009; Myosin_N.
```

```
DR InterPro; IPR002928; Myosin_tail.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR PRODOM; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS00096; IQ; 1.
SQ SEQUENCE 1940 AA; 223318 MW; EE2642E1E29FDDC2 CRC64;

Query Match 13.7%; Score 425; DB 6; Length 1940;
Best Local Similarity 23.1%; Pred. NO. 6.9e-08;
Matches 169; Conservative 143; Mismatches 25; Indels 162; Gaps 23;

Qy 4 LSLELMKLRNKRRETKMRSMVMYQEGMELKLQATQDLTESKGIQVQLEKGLYSIEKEID 63
Db 837 LFFRIKPLKLSAETEKEMATKKE-----FORKTDELAKSEAKRKELEKMWTLLEKND 891
Qy 64 EKCETEKLLYIQETISCASDOVEKCV-----DIAOLEEDLKEKDRILSLKQSLLEN 116
Db 892 LOLQVQSAEGLADAEERCDOLIKTQLEAKIKEVTERAEDEEELNAELTAKRKKLEDE 951
Qy 117 IT-FSKQIEDLTVKCOLLETERDNVSKDRERAETLSAEMOILTERLALERQVEYKLOQK 175
Db 952 CSELKDDIDDLLETLAKVEKEKHATENK-----VKNLTEEMAGLDETIAKLTREKKAQEA 1007
Qy 176 ELQOSLILQOEKE-----LSARLOOQLCSFOEEMTSEKNV-----FRKELKLA 218
Db 1008 HQQTLDLDDLOAEEDKVTNTLTAKTKLEQVDDLEGSLEQEKLRMDLERAKRKKLEGLDKLA 1067
Qy 219 LAELDVAQOKEEBSERLVKOLE-----EERKSTAEQOLTRLDNLLREKEVELEKHI 268
Db 1068 QESINDIENEKQDLDEKLKKEFEISNLQSKIEDQALGILQKIKELQARIELEEEI 1127
Qy 269 AAHAQAIIIAQEKYNDTAQSLRDVTAQLESVOEKYNDTAQSLRDVTAQLESQEKYNDTA 328
Db 1128 EAERASRAKAEKQSDLSRELEIEISERLE-----EAGGATSQIEMNKKREAEFQ 1177
Qy 329 QSLRDVTAQLESQEKYNDTAQSLRDVTAQLESVOEKYNDTAQSLRDVSAQLESYKSSTL 388
Db 1178 KMRRD-----LEPATLQHEATAAALRKKHA--DSVAE-LGEQIDNLRVKQKLEKEKSEM 1230
Qy 389 KEIEDLKLNTLTQEKVMAEKSVEDVQOQILTAESTNQEYARVMVQDL-ONRSTLKEE-- 445
Db 1231 MEIDDLASNVETISKAGNLEKMCRTLEDQVNELSKSEEEQORLINDLTQGRQLQTESG 1290
Qy 446 -----BIKEITSSFEKITDLKNLROQ-----DED-FRK 474
Db 1291 EFSRQLEKEALVSOLSRGKQAFQOIEELKQLEEEIKAKNALAHGLQASRHDCDLLRE 1350
Qy 475 QLEEKGRKTAENVMTELTMEINKRLLY-----EELYEKTQPFQOQLD----- 520
Db 1351 QYEEQESKAELQRLALSKANTEVAQWRTKYETDAIQRTEELEEAQKLAQRLQDAEERHVE 1410
Qy 521 -----FEAEKQALLNE-----HGATQEOQLN-----KIRDSYA 547
Db 1411 AVNAKASLEKTKQRLQNEVEDLMDVETNNAACAALDKKQNFDKILAENKHKRYEETHA 1470
Qy 548 QI-----IGHONLKQK-----IKHVVKLKDENSOLKSEVSKLRSQLVK--RKQNE 590
Db 1471 ELEAAQKEARSLSGTELFKMKNAYESLDQLETLRKKNLQOEISDLTEQTAEGGKRMHE 1530
Qy 591 L---RLQELDKA 600
Db 1531 LEKIRKQVEQEK 1543

RESULT 9
Q9TV63 ID Q9TV63 PRELIMINARY; PRT; 1939 AA.
AC Q9TV63;
```

```
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Myosin heavy chain 2a.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LANDRACE; TISSUE=SKELETAL MUSCLE;
RA Chikuni K., Tanabe R., Muroya S., Nakajima I.;
RT "Differences in molecular structure among the porcine myosin heavy
RL chain-2a, -2x, and -2b isoforms.";
DR Meat Sci. 57:311-317(2001).
DR EMBL; AB025260; BAA82144.1; -.
DR HSSP; P13538; 2MYS.
DR InterPro; IPR000048; IQ_region..
DR InterPro; IPR001609; myosin_head.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR PRODOM; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1..
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 1.
DR SQ SEQUENCE 1939 AA; 223149 MW; B520182094179343 CRC64;

Query Match
Best Local Similarity 13.6%; Score 423; DB 6; Length 1939;
Matches 170; Conservative 147; Mismatches 254; Indels 168; Gaps 23;

QY 9 MKLRNKRFTKMSMMVKOE--GMELKQATQKDLTSEKGIQVLEGKLVSEIEKEIDEKC 66
DB 834 MKLFKIKPLKLSAESEKEMANMKEEFKTKDELAKSEAKRKELEKMMVTLKENDLQL 893
QY 67 ETEKLELYQIEISCASDQVEKCV-----DIALEEDLKEKDRILSLKOSLEENIT- 118
DB 894 QVQAEADALAEERCQDLIKTKIQLEAKIKEVTERAEDEEINAEIATAKRRKLEDECS 953
QY 119 FSKQIEDLVTKQQLLETDRNLVSKDRERAETLSAEMQILTERLALEROEYKELQKQELQ 178
DB 954 LKKDIDDLTLAKYKEKEHATENK----VKNLTEEMAGLDIETIAKTKKALQEAHQ 1009
QY 179 SOSLQOQKE-----LSARLQOQCSFOEEMTSEKNV-----FKDELKLAEE 221
DB 1010 TLDDLQAEEDKYNVTLTKAKTKLEQVDDLEGSLEQEKLRMDLERAKRKLGLDKLAQES 1069
QY 222 LDVQOQKEQSRIVKOLB-----EERKSTAEQILTRLDNLLREKEVELEKHIAH 271
DB 1070 IMDIENEKQOLDEKQKKEFEFETSNLSQKIEDQALQIKKIQELQARIELEEEIEAE 1129
QY 272 AQAILIAQEKYNDTAQSLRDVTAQLESQVEKYNDTAQSLRDVTAQLESQVEKYNDTAQSL 331
DB 1130 RASRAKAEQRSDLSRELEIEISERLE-----EAGGATSAQIEMNKKREAFQKWR 1179
QY 332 RDVTAQLESQVEKYNDTAQSLRDVTAQLESQVEKYNDTAQSLRDVTAQSLRDVTAQSL 391
DB 1180 RD-----LEEATLQHEATAALRKHIA--DSVAE-LGEQIDNLRQVKQKLEKEKSEMKMEI 1232
QY 392 EDLKLLENLTLOEKVMAEKSVEDVQOQILTAESTNQEYARVQDL-QNRSTLKEB----- 445
DB 1233 DLASNMTVSKAGNLKRCMTLEDQSLSKSEEEQORLINDTAQGRGLQTSGBFS 1292
QY 446 -----BIKETTSFSEKIKITDLKNQLRQ-----DED-FRKOLE 477
DB 1293 RQDKEKALVSQLSRGQAYTQIQIEELKRLQEEIETAKNALAHALQSSRHCDLLREQYE 1352
QY 478 EKGKRTAEKENVMTELTMWINKWRLLY-----EELYETKTPFQQQLDA----- 520
```

```
DB 1353 EQESKAEQLRALSKANTEVAQWRTKYETDAIQTRELEEAQKLAQRLQAAAEHVEAVN 1412
QY 521 -----FEAEKQALLNE-----HGATQEQQLN-----KIRDSVAQL- 549
DB 1413 AKCASLETKTKORLQNEVEDMLDVERTNACAALDKORNFDKIIAEWKQRYEETHAELE 1472
QY 550 -----LGHONLKQK-----IKHVVKLKQDENSQKSEVSKLSQKL----- 583
DB 1473 ASQKEARSLSGLTELFKMKNAYEESLQOLETFLKRENKNLQOEISDLTEQTAEGGKRIHELEK 1532
QY 584 VKR--KQNELRLQGBELDKA 600
DB 1533 IKKQVEQEKSEIQAALAEA 1551

RESULT 10
Q9DGM5 PRELIMINARY; PRT; 1944 AA.
ID Q9DGM5;
AC Q9DGM5;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Fast myosin heavy chain isoform 2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang Q., Bandman E.;
RT "Seven skeletal myosin heavy chain genes (MyHC) are organized as a
RT multigene complex in the chicken genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF272033; AAF99314.1; -.
DR HSSP; P13538; 2MYS.
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR001609; myosin_head.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR InterPro; IPR000533; Tropomyosin.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR PRINTS; PR00194; TROPOMYOSIN.
DR PRODOM; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR SQ SEQUENCE 1944 AA; 223211 MW; 154F438220072D68 CRC64;
```

Query Match 13.5%; Score 418; DB 13; Length 1944;
Best Local Similarity 22.8%; Pred. No. 1.2e-07;
Matches 172; Conservative 139; Mismatches 244; Indels 200; Gaps 26;

```
QY 9 MKLRNKRFTKMSMMVKOE--GMELKQATQKDLTSEKGIQVLEGKLVSEIEKEIDEKC 66
DB 838 MKLSFKIPLKLSAESEKEMANMKEEFKTKELAKSAKRKELEKMMVTLKQENDLQL 897
QY 67 ETEKLELYQIEISCASDQVEKCV-----DIALEEDLKEKDRILSLKOSLEENIT- 118
DB 898 QVQAEADALAEERCQDLIKTKIQLEAKIKEVTERAEDEEINAEIATAKRRKLEDECS 957
QY 119 FSKQIEDLVTKQQLLETDRNLVSKDRERAETLSAEMQILTERLALEROEYKELQKQELQ 178
DB 958 LKKDIDDLTLAKYKEKEHATENK----VKNLTEEMAGLDIETIAKTKKALQEAHQ 1013
QY 179 SOSLQOQKE-----LSARLQOQCSFOEEMTSEKNV----- 210
DB 1014 TLDDLQAEEDKYNVTLTKAKTKLEQVDDLEGSLEQEKLRMDLERAKRKLGLDKLMSQDT 1073
```

```
Qy 211 -----FKEELKIALAELDAVOQKEQSERLVKOLEEERKSTAEQTLRLDNLREK 260
Db 1074 IMDLENDKQOLDEKLKKDFEISQISQIEDEQALGMQLOKKIK--ELQARIEEL--EE 1128
Qy 261 EVELEKHIAHAQAAILIAQEKYN--DTAQSRLDVTVAQLESVQEKYNDNTAQSRLDVTVAQLES 319
Db 1129 EIEAERTSRKA-----EKHRADLSRELEFEISERLE---EAGGATATQI--DMNKKREA 1177
Qy 320 EOEKYNDTAQSRLDVTVAQLESQEKYNDNTAQSRLR---DVTVAQLESVQEKYNDNTAQSRLD 375
Db 1178 EFQKMR---RDLEATLQHEA-----TAAALRKKHADSTAEL-----GEQIDNLQR 1220
Qy 376 VSAQLESYKSTLKEIEDLKENLTLOEKVAMAEKSVEDVQQOILTAESTNQEYARMVQD 435
Db 1221 VKOLEKESKELKWEIDDLASNMESVS KAKANLEKMCRTLEDQLSKIKSKEEHRQMIND 1280
Qy 436 LQ-NRSTLKEE-----EIKETSSFLEKITDNLKQLRQ----- 468
Db 1281 LSTQARLQTESGYSRQVEEKDALISQLSRGKQAFQTQIEBELKRLHEEIKAKNALAHA 1340
Qy 469 -----DED-FRKOLEEKGKRTAEKNVMTLWINKWRLLY-----EELYEKTTP 513
Db 1341 LQSAHDCDLLREQYEEQEKAGELQALSKANSEVAQWRTKYETDALTQRTTEELEBAKK 1400
Qy 514 FQOQLD-----AFAEAKQALLNE-----HGATQEQNLKIRDSYAQL 549
Db 1401 LAQRLQDAEEHVEAVNAKASLEKTKQRLQNEVEDLMIDVERANACARLDKKQKQKFDKI 1460
Qy 550 LGHONLK-----OKKHVVYKLDENSQLSKSEVSKL 579
Db 1461 LAEWKQKYETQAELEASQESRSLSLSTELFKMKNAYEESLDHLETLKRENKLNQOEISDL 1520
Qy 580 RSOL-----VKR--KQNELRLQGLDKA 600
Db 1521 TEQIAGGKAHLEKVKQKEQEKSEIQAALAEA 1555

RESULT 11
Q9BE40 PRELIMINARY; PRT; 1938 AA.
AC Q9BE40;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE MYOSIN heavy chain 2x.
GN MYHC-2X.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HOLSTEIN; TISSUE-SKELETAL MUSCLE;
RA Chikuni K., Muroya S., Nakajima I.;
RT "Sequencing of the bovine myosin heavy chain isoforms.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB059399; BAB40921.1;
DR HSSP; P13538; 2MIS.
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR001609; myosin_head.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 1.
SQ SEQUENCE 1938 AA; 222988 MW; 95D05BE1B8978D40 CRC64;
```

Query Match

13.4%; Score 417; DB 6; Length 1938;

Best Local Similarity 23.0%; Pred. No. 1.3e-07;

Matches 170; Conservative 147; Mismatches 254; Indels 168; Gaps 24;

```
Qy 9 MKLNRKRETKMRSMVMKQOE--GMELKQATQKDTESKGKIVQLEKGLVSIKEKIDKCK 66
Db 833 MKLYFKIKPLLKSAETEMANMKEFEKTEELAKSEAKRKELEKVMVTITQEKNDLQL 892
Qy 67 ETEKLEVIQIISCASDOVEKCV-----DIAQLEEDLKEKOREIILSLKOSLENT- 118
Db 893 QVQSEADALADAERCOQLIKTIQLEAKIKEVTERAEDEEIEINAEIAKAKRKEDECE 952
Qy 119 FSKQIEDLTVCQQLTERDNLVSKDRERAETLSAEMOILTERLALEREQVEKLOKQELQ 178
Db 953 LKKDIDDLLETLAKVEKHAENK---VKNLTTEAGLDETIAKLTKEKKALQEAHQ 1008
Qy 179 SQSLQOQKE-----LSARLQOOLCSFQEMTSEKNV-----FKELKAL-A 220
Db 1009 TLDDLQAEEDKVNTLTAKTKLEQVDDLEGSLEQEKLRMDLERAKRKLKGLDLAOGS 1068
Qy 221 ELDAVOQKEEESERLVKQ-----LE---EERKSTAEQTLRLDNLREKEVELEKHIAH 271
Db 1069 TMDIENDKQQLDEKLKKKEFEMSNIQSKIEDEQALAMQKKIKELQARIIELEEEIEAE 1128
Qy 272 AQAILIAQEKYNDTAQSRLDVTVAQLESVQEKYNDNTAQSRLDVTVAQLESQEKYNDTAQSL 331
Db 1129 RASRAKAEKQKSDLSRELEFEISERLE-----EAGGATSAQIEMNKKREAFQKMR 1178
Qy 332 RDVTAQLESQEKYNDTAQSRLDVTVAQLESVQEKYNDNTAQSRLDVSQALESYKSTLKEI 391
Db 1179 RD----LEEATLQHEATAAALRKKHA--DSVAE-LGEQIDNLQVKKQLEKSEKSMKEI 1231
Qy 392 EDLKENLTLOEKVAMAEKSVEDVQQOILTAESTNQEYARMVQDQNLQSTLKEEIKET 451
Db 1232 DLLASNMETVSKAGNKLEKMCRALEDQLSELKTEDEQOQLINDLTQARLQTESGFS 1291
Qy 452 -----SSFLEKITDNLKQLRQ-----DED-FRKOLE 477
Db 1292 RQLEKDALVSQLSRGKQAFQTQIEELARQLEEEIKAKSALAHALQSAHDCDLLREQYE 1351
Qy 478 EKGKRTAEKENVMTLWINKWRLLY-----EELYEKTTPFQOQLD----- 519
Db 1352 EEQEGKAEQLQMSKANSEVAQWRTKYETDAIQRTEEELEAKKLAQRLQDAEEHVEAVN 1411
Qy 520 ----AFAEAKQALLNE-----HGATQEQNL-----KIRDSYAQL 550
Db 1412 AKCASLEKTKQRLQNEVEDLMIDVERTNAACAALDKKORNFDKILSEWKQKYETHAELE 1471
Qy 551 GHQNLKQ-----KIKHWK-----LKDENSQLSKSEVSKLSQL----- 583
Db 1472 ASQKESRSLSLSTELFKMKNAYEESLDQLETLKRENKLNQOEISDLTEQIAGGKRIHELEK 1531
Qy 584 VKR--KQNELRLQGLDKA 600
Db 1532 VKKQVEQEKSEIQAALAEA 1555

RESULT 12
Q14221 PRELIMINARY; PRT; 1410 AA.
AC Q14221;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Endosome-associated protein.
GN EEA1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
```

RX MEDLINE-95286647; PubMed-7768953;
RA Mu F.T., Callaghan J.M., Steele-Mortimer O., Stenmark H., Parton R.G.,
RA Campbell P.L., McCluskey J., Yeo J.P., Tock E.P., Toh B.H.;
RT "EEA1, an early endosome-associated protein. EEA1 is a conserved
RT alpha-helical peripheral membrane protein flanked by cysteine
RT 'fingers' and contains a calmodulin-binding IQ motif.";
RL J. Biol. Chem. 270:13503-13511(1995).
DR EMBL; I40157; AAA79121.1; .
DR InterPro: IPR000822; Znf_C2H2.
DR InterPro: IPR000306; Znf_FYVE.
DR Pfam; PF01363; FYVE; 1.
DR SMART; SM00064; FYVE; 1.
DR SMART; SM00355; Znf_C2H2; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 1.
KW DNA-binding; Zinc-finger.
SQ SEQUENCE 1410 AA; 162496 MW; DF0F9464D70A8AED CRC64;

Query Match 13.4%; Score 415; DB 4; Length 1410;
Best Local Similarity 22.8%; Pred. NO. 1.2e-07;
Matches 168; Conservative 153; Mismatches 223; Indels 192; Gaps 28;

QY 1 MRALSLEMLKLN-----KRETKMRSMVYKQEG---MELK-----LQAT--OK 38
DB 283 VRVYVOELQKLKSSVNETQKNQTLTENLLKKEQDYTKLEEKHNESVSKKNIQATLHQK 342
QY 39 DLTESGKGIVQLEGKLVSTKE-----KIDEKCE-TEKLLEYIOEISCASDQVEKCKVD 91
DB 343 DL-----DCQQLQSLRLSASETSLSRIHVSELSKGEATQKLKEELSEV---ETKYQHLKAE 394
QY 92 IAQLEDLKEKREILSKQLEENITFSKOIEDLVTKVQCLLETDRNLVSKDRERAETLSAE 151
DB 395 FKLOLQOQREKEQHGLOLQSEINQ-----LHSLKLETEROLGAEHGR-----L 437
QY 152 SAEMQILTERLALEREYKELQKELQSQSLLOQKELSAKRLQOQL-----CSFQ 201
DB 438 KEQRLSSEKLMKEQOVADLQKLRLSEELKEKVTNSTELQHLQDKTKQOQHQQALQ 497
QY 202 EEMTEKKNVFEELKALAEADVQKQESERLVKQ-----LEERK----- 244
DB 498 QSTAKLRAQNDLEQVLRLQICEKQKQINLEALLQKSKENISLLEKREDLYAKIQAGE 557
QY 245 -----STAQQLTRDLNLLREKEVELEKHAHAQAAILIAQEKYNDTAQSLRDVTAQLESVQE 301
DB 558 GETAVLNQLOEKNHNTLQDQVTLQTEKLNQSESHKQAEHLQDVQEQK---AHLRAAQD 614
QY 302 KYNDTAQSLRDVTAQLESQEKYND-----TAQSLRDVTAQLESQ 342
DB 615 RVLSLETSVNELNSQLNESKEKVSQLDIQIKAKTELLISAEAKTAQR-ADLQNLHDTAQ 673
QY 343 EKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVSAQLESYKSTL---KEIEDL----- 394
DB 674 NALQDKHQLNKITTLQDQVTAQKQKQHCQSLESHLKEYKEYLSLEQKTEELGQIK 733
QY 395 KLENLTDLNLLREKEVELEKHAHAQAAILIAQEKYNDTAQSLRDVTAQLESVQEKYNDTA 451
DB 734 KLEADSLVK-ASKEQALQDLQOQ-----RLNTDLELRATLSKQLEMEKEIV 781
QY 452 SSF-----LEKITYDLKNLRQODEDFRKQLEKGGKRTAEKNVMTLMTWINKRWLLY 504
DB 782 STRLDLQKKSALRESIKQKLTQKEE--KQILKQDFETLSQET-----KIQH 827
QY 505 BELYEKTFFQOQLDAFAEAKALLNEHCATQEQNLNKIRDSY-----AQL 549
DB 828 EELNRIQTVTELQKVKMEKALMTSTLVKDKLSKVSQSLKNSKSEFEKENQKGA 887
QY 550 LGHQNLKQIKRHVVKLDKENS-----QLKSVSKLRSOLVKR----- 586
DB 888 LDLEXTKELKHQLOVQMENTLKEQELKKSLEKEKASHQLKLENSMQEQLIAQNTL 947
QY 587 KQN---ELRLQELDK 599
DB 948 KQNEKEEQLOQGNINE 963

RESULT 13

Q15075 PRELIMINARY; PRT: 1411 AA.
ID Q15075
AC Q15075;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Endosomal protein.
GN P162.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Seelig H.P.;
RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; X78998; CAA55632.1; .
DR InterPro: IPR000822; Znf_C2H2.
DR InterPro: IPR000306; Znf_FYVE.
DR Pfam; PF01363; FYVE; 1.
DR SMART; SM00064; FYVE; 1.
DR SMART; SM00355; Znf_C2H2; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 1.
KW DNA-binding; Zinc-finger.
SQ SEQUENCE 1411 AA; 162465 MW; C3B17777FE34B6BD CRC64;

Query Match 13.3%; Score 414; DB 4; Length 1411;
Best Local Similarity 23.0%; Pred. NO. 1.3e-07;
Matches 168; Conservative 150; Mismatches 220; Indels 192; Gaps 28;

QY 7 ELMKLN-----KRETKMRSMVYKQEG---MELK-----LQAT--OKDLTSEK 44
DB 289 ELQKLKSSVNETQKNQTLTENLLKKEQDYTKLEEKHNESVSKKNIQATLHQKDL---- 344
QY 45 GRIVOLEKGLVSTKE-----KIDEKCE-TEKLLEYIOEISCASDQVEKCKVDIAQLEE 97
DB 345 -DCQQLQSLRLSASETSLSRIHVSELSKGEATQKLKEELSEV---ETKYQHLKAEFKQLOQ 400
QY 98 DLKEXDRILSKQLEENITFSKOIEDLVTKVQCLLETDRNLVSKDRERAETLSAEHQI 157
DB 401 QREEKEQHGLOLQSEINQ-----LHSLKLETEROLGAEHGR-----LKEQRL 443
QY 158 LTERLALEREYKELQKELQSQSLLOQKELSAKRLQOQL-----CSFQEMTSE 207
DB 444 SSEKLMDEKQOVADLQKLRLSEELKEKVTNSTELQHLQDKTKQOQHQQALQOQSTAK 503
QY 208 KNVFEELKALAEADVQKQESERLVKQ-----LEERK-----STA 247
DB 504 LREAQNDLEQVLRLQIGDKDQKQINLEALLQKSKENISLLEKREDLYAKIQAGEGTA 563
QY 248 EQLTRDLNLLREKEVELEKHAHAQAAILIAQEKYNDTAQSLRDVTAQLESVQEKYNDTA 307
DB 564 NQLOEKNHNTLQDQVTLQTEKLNQSESHKQAEHLQDVQEQK---AHLRAAQDRVLSLE 620
QY 308 QSLRDVTAQLESQEKYND-----TAQSLRDVTAQLESQEKYNDT 348
DB 621 TSVNELNSQLNESKEKVSQLDIQIKAKTELLISAEAKTAQR-ADLQNLHDTAQNALQDK 679
QY 349 AQSLRDVTAQLESVQEKYNDTAQSLRDVSAQLESYKSTL---KEIEDL-----KLENLT 400
DB 680 QOELNKITTLQDQVTAQKQKQHCQSLESHLKEYKEYLSLEQKTEELGQIKLEADS 739
QY 401 LQEKVAMAEKSVEDVQOQILTAESTNQEYARMVQDLQNRSLTKEEBI---KEITSSF--- 454
DB 740 LEVK-ASKEQALQDLQOQ-----RLNTDLELRATLSKQLEMEKEIVSTRLD 787
QY 455 ----LEKITYDLKNLRQODEDFRKQLEKGGKRTAEKNVMTLMTWINKRWLLYELYEK 510
DB 788 LQKKSEALRESIKQKLTQKEE--KQILKQDFETLSQET-----KIQHELNNR 833


```

SQ SEQUENCE 1939 AA; 2231172 MW; B702ADB599602ECB CRC64;

Query Match 13.2%; Score 411; DB 6; Length 1939;
Best Local Similarity 22.7%; Pred. No. 2.2e-07;
Matches 168; Conservative 151; Mismatches 252; Indels 168; Gaps 25;

QY 9 MKLNRKRETKMRSMVMKQF--GMELKQATQKDLTESKGVKIQVLEGKLVSTEEKIDKCK 66
   ||| ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
Db 834 MKLYFKIKPLKLSAETEKEMANKEEFKTKSLAKAEKAKKLEKFMVAKQEKNDLQL 893
   : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
QY 67 ETEKLLEYIQIBISCASDOVEKCV-----DIAQLEDLDEKDEILSLKQSLSENIIT- 118
   : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
Db 894 QVQAEADSLADAEBCDOLIKTKIQLEAKIKEVTERAEDEEINAEINAKKLEDECESE 953
   : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
QY 119 FSKQIEDLTVKQCQLLETERDNVLVSKDRERATLSAEMOILTERIALERQYEKLOQKELQ 178
   ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
Db 954 LKKDIDDELTLAKVEKHEKATENK---VKNLTEEMAGLVDETIAKTKKXKALQEAHQ 1009
   : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
QY 179 QSLSLQOEKE-----LSARLQOLCSFQEEMTSEKNV-----FREEKLKIAL-A 220
   ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
Db 1010 TLDLQAEEDKVNLTFRAKTKLEQVDDLEGSFQEKRLMDLRKRLKLEGLDKLAQES 1069
   : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
QY 221 ELDAVQOKEEQSERLVKQ-----LE---EBRKSTAEQLTRLDNLLRKEVEKHKHTAAH 271
   : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
Db 1070 TMDIENDKQQLDEKLKKKFEFMSNLQSKIEDEQALAMQLOKKIKELQARIEELEEETAE 1129
   : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
QY 272 QAATLIAQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESQEKYNDTAQSL 331
   : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
Db 1130 RASRAKAEKORSLSRELEESERLE-----EAGGATSAQIEMNKKREAEFOKMR 1179
   : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
QY 332 RDVTAQLESQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVSAQLESYKSTSLKEI 391
   ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
Db 1180 RD-----LEEATLQHEATATLRKHA--DSVAE--LGEQIDNLRVQKLEKEKSEMKEI 1232
   : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
QY 392 EDZKLENLTQEKVAMAEKSVEDVQOQILTAESTNQYARMVDOL-QNRSTLKEE----- 445
   : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
Db 1233 DDLASNMETVSKAGNLEKMCRTLEDQLSLTKEEEQORLINDLTQAORLQTESGEYS 1292
   : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
QY 446 -----EIKEITTSFLEKTLDLKNOLRQ-----DED-FRQOLE 477
   : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
Db 1293 ROLDEKDTLVSQLSRGKQATFQQLTEELKQLEEEIKAKALAHAVQSRHDCDLLRQOYE 1352
   : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
QY 478 EKGKRTAEKNVMVTELTMEINKWRLLY-----EELYEKTKPFOQLD----- 519
   : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
Db 1353 EEQAKAELQAMSKANSEVAQWRTKVETDAIQTEELEAKKLAQRLQDAEEHVAVN 1412
   : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
QY 520 ----AFAEKQALLNE-----HGATQEQOLN-----KIRDSYAQLL 550
   : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
Db 1413 AKCASLEKTKORLQONEVEDLMIDVERSNAACALDKKORNFDKTLAEMKQKYETHAELE 1472
   : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
QY 551 GHQNLQ-----KIKHVVK-----LKDENSOLKSEVSKRSULS----- 583
   : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
Db 1473 ASQKESRSLSTELFKVNAYEESLDQLETLKREKNLNQOEISDLTQIABGGKKRIHELEK 1532
   : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
QY 584 VKR--KONELRLRGELDKA 600
   : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
Db 1533 IKQVEQEKSEIQAALEEA 1551
   : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :

```

Search completed: December 20, 2002, 15:14:05
Job time : 46.6431 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:09:23 ; Search time 17.2176 Seconds
(without alignments)
3523.197 Million cell updates/sec

Title: US-09-685-010-48
Perfect score: 3104
Sequence: 1 MRALSLEMLKLRNKRRETKMR.....FCHASKENFTPLKEGPNCC 631
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3098	99.8	631	2 JC4298	hyaluronan recepto
2	1953	62.9	725	1 JC5016	hyaluronan recepto
3	431.5	13.9	1790	2 S67593	transport protein
4	421	13.6	1938	2 A59293	skeletal myosin he
5	415	13.4	1410	1 A57013	early endosome ant
6	415	13.4	1940	1 A24922	myosin heavy chain
7	411	13.2	1938	1 JX0178	myosin heavy chain
8	408.5	13.2	1690	2 T13030	microtubule bindin
9	408	13.1	1130	2 T34081	hypothetical prote
10	404	13.0	1940	2 A29320	myosin heavy chain
11	397	12.8	1940	1 S04090	myosin heavy chain
12	393	12.7	1939	2 T48175	myosin heavy chain
13	391	12.6	2954	2 T14156	kinesin-related pr
14	390	12.6	1738	2 T14867	interaptin - slime
15	389.5	12.5	1935	1 A37102	myosin beta heavy
16	387	12.5	1938	1 S06005	myosin alpha heavy
17	387	12.5	1938	2 T49464	alpha cardiac myos
18	386	12.4	1392	2 A43336	microtubule-vesicl
19	384.5	12.4	2139	2 T18296	myosin heavy chain
20	384	12.4	1979	2 C71622	hypothetical prote
21	383	12.3	1427	2 S22895	restin - human
22	380.5	12.3	1931	2 A59234	slow myosin heavy
23	379.5	12.2	1937	2 T38055	myosin heavy chain
24	379.5	12.2	1938	1 A40997	myosin heavy chain
25	378.5	12.2	1937	2 T48153	myosin heavy chain
26	378	12.2	1837	2 T41023	probable nuclear p
27	377	12.1	1935	1 S06006	myosin beta heavy
28	376.5	12.1	1978	2 A70387	conserved hypothet
29	376.5	12.1	1999	1 S21801	myosin heavy chain

ALIGNMENTS

RESULT 1

JC4298
hyaluronan receptor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 16-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 01-Dec-2000
C:Accession: JC4298; A42925; A41923; S21586
R:Entwistle, J.; Zhang, S.; Yang, B.; Wong, C.; Li, Q.; Hall, C.L.; A, J.; Mowat, M.;
- Gene 163, 233-238, 1995
A:Title: Characterization of the murine gene encoding the hyaluronan receptor RHAMM.
A:Reference number: JC4298; MUID:96011639; PMID:7590272
A:Accession: JC4298
A:Molecule type: mRNA
A:Residues: 1-631 <ENT>
A:Cross-references: EMBL:X64550
A:Experimental source: 3T3 fibroblast
R:Hardwick, C.
- J. Cell Biol. 118, 753, 1992
A:Reference number: A42925; MUID:92348516; PMID:1639856
A:Contents: erratum
A:Accession: A42925
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 155-376, 'S', 378-504, 'E', 506-631 <HAR>
A:Cross-references: GB:X64550
A:Note: authors translated the codon AGT for residue 377 as Thr and CTA for residue 5
R:Hardwick, C.; Hoare, K.; Owens, R.; Hohn, H.P.; Hook, M.; Moore, D.; Cripps, V.; Au
- J. Cell Biol. 117, 1343-1350, 1992
A:Title: Molecular cloning of a novel hyaluronan receptor that mediates tumor cell mo
A:Reference number: A41923; MUID:92299690; PMID:1376732
A:Accession: A41923
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 155-376, 'S', 378-504, 'E', 506-507, 'L', 508-630 <HA2>
A:Cross-references: GB:X64550
A:Note: this sequence has been corrected in reference A42925
C:Comment: This protein regulates cell motility and transformation, and focal adhesio
C:Genetics:
A:Gene: rhamm
A:Introns: 20/3; 54/2; 79/2; 137/3; 187/3; 259/2; 382/2; 431/2; 482/2; 515/3; 574/3;
A:Superfamily: hyaluronan receptor
C:Keywords: glycoprotein; receptor
F:260-382/Region: 21 residue repeats
F:516-574/Region: hyaluronan binding #status predicted
F:575-625/Region: hyaluronan binding #status predicted
F:116,283,304,325,346,367,398,438,619/Binding site: carbohydrate (Asn) (covalent) #st
Query Match 99.8%; Score 3098; DB 2; Length 631;
Best Local Similarity 99.7%; Pred No. 3.3e-100;
Matches 629; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Oy 1 MRALSLEMLKLRNKRRETKMRSMVKQEGMELKQATKDLTESKIVQLEGKLVSIKKE 60
|||||

Db 1 MRALSLEMLKLNKRETKRMVMVQEGMELKQATQKDLTSGKIVQLEGKLSIEKE 60
QY 61 KIDEKCEKLELEYIQEISCASDQVEKCKVDIAQLEEDLKEKDREILSLKQSLSEENIT-F 120
Db 61 KIDEKCEKLELEYIQEISCASDQVEKCKVDIAQLEEDLKEKDREILSLKQSLSEENIT-F 120
QY 121 KOIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALEROEYKIQKQKELQSO 180
Db 121 KOIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALEROEYKIQKQKELQSO 180
QY 181 SLLOQKELSLRLOOQLCSFOEEMTSKKNVFEELKLAELDAVQOKEQSERLVKQLE 240
Db 181 SLLOQKELSLRLOOQLCSFOEEMTSKKNVFEELKLAELDAVQOKEQSERLVKQLE 240
QY 241 EERKSTAEOLTRLDNLLREKEVELEKHAHAQAAILIAQEKYNDTAQSLRDVTAQLESVQ 300
Db 241 EERKSTAEOLTRLDNLLREKEVELEKHAHAQAAILIAQEKYNDTAQSLRDVTAQLESVQ 300
QY 301 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 360
Db 301 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 360
QY 361 SVQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 420
Db 361 SVQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 420
QY 421 TAESTNOEYARMVQDLQNRSTLKEEIEKITSFLEKITDKNLQROODEDFRKOLEEK 480
Db 421 TAESTNOEYARMVQDLQNRSTLKEEIEKITSFLEKITDKNLQROODEDFRKOLEEK 480
QY 481 KRTAEKENVMTLMEINKWRLLYELYEKTPFQOQDAPFAEAKOALLNEHGATQEQLN 540
Db 481 KRTAEKENVMTLMEINKWRLLYELYEKTPFQOQDAPFAEAKOALLNEHGATQEQLN 540
QY 541 KIRDSYAQLLGHQNLKQKIKHVVKLDKNSQLKSEVSKLRSOLVRKQNELRLQGELODKA 600
Db 541 KIRDSYAQLLGHQNLKQKIKHVVKLDKNSQLKSEVSKLRSOLVRKQNELRLQGELODKA 600
QY 601 LGIRHFDPSKAFCHASKENFTPLKEGPNCC 631
Db 601 LGIRHFDPSKAFCHASKENFTPLKEGPNCC 631

RESULT 2
JC5016
hyaluronan receptor - human
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 29-Sep-1999
C:Accession: JC5016
R:Wang, C.; Entwistle, J.; Hou, G.; Li, Q.; Turley, E.A.
Gene 174, 299-306, 1996
A:Title: The characterization of a human RHAMM cDNA: Conservation of the hyaluronan-binding site in the human and mouse
A:Reference number: JC5016; MUID:97045829; PMID:8890751
A:Contents: breast
A:Accession: JC5016
A:Molecule type: mRNA
A:Residues: 1-725 <W>
A:Cross-references: GB:U29343
A:Note: it is uncertain whether Met-1 or Met-196 is the initiator
C:Comment: This receptor regulates focal adhesion turnover, and regulates human breast cancer cell growth
C:Genetics:
A:Gene: GDB:HMMR; RHAMM
A:Cross-references: GDB:683209; OMIM:600936
A:Map position: 5q33.2-5pter
C:Superfamily: hyaluronan receptor
Query Match 62.9%; Score 1953; DB 1; Length 725;
Best Local Similarity 64.8%; Pred. No. 9.7e-61;
Matches 411; Conservative 49; Mismatches 86; Indels 88; Gaps 3;

QY 1 MRALSLEMLKLNKRETKRMVMVQEGMELKQATQKDLTSGKIVQLEGKLSIEKE 60
Db 164 LRLSLEMLKLNKRETKRMVMVQEGMELKQATQKDLTSGKIVQLEGKLSIEKE 223

QY 61 KIDEKCEKLELEYIQEISCASDQVEKCKVDIAQLEEDLKEKDREILSLKQSLSEENIT-F 119
Db 224 KIDEKCEKLELEYIQEISCASDQVEKCKVDIAQLEEDLKEKDREILSLKQSLSEENIT-F 283
QY 120 SKOIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALEROEYKIQKQKELQSO 179
Db 284 SKOVEDLNVKCOLLEKEKEDHVRNRHNEHNLNAEMQNLKQKFILOEQEHEKLOQKELQI 343
QY 180 QSLLOQKELSLRLOOQLCSFOEEMTSKKNVFEELKLAELDAVQOKEQSERLVKQLE 239
Db 344 DSLLOQKELSLRLOOQLCSFOEEMTSKKNVFEELKLAELDAVQOKEQSERLVKQLE 403
QY 240 EERKSTAEOLTRLDNLLREKEVELEKHAHAQAAILIAQEKYNDTAQSLRDVTAQLESV 299
Db 404 EEEAKRAEELKLEKLGKAELEKSSAAHTQATLLLOEKYDNVQSLQSLDVT----- 457
QY 300 QEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 359
Db 458 ----- 457
QY 360 ESVQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 419
Db 458 -----AQFESYKALTASEIEDLKLENSLQEKAAKNAEDVQHOI 499
QY 420 LTAESTNOEYARMVQDLQNRSTLKEEIEKITSFLEKITDKNLQROODEDFRKOLEEK 479
Db 500 LATESNOEYARMVQDLQNRSTLKEEIEKITSFLEKITDKNLQROODEDFRKOLEEK 559
QY 480 KRTAEKENVMTLMEINKWRLLYELYEKTPFQOQDAPFAEAKOALLNEHGATQEQLO 539
Db 560 EGRKAENKTAETLMEINKWRLLYELYEKTPFQOQDAPFAEAKOALLNEHGATQEQLO 619
QY 540 NKIRDSYAQLLGHQNLKQKIKHVVKLDKNSQLKSEVSKLRSOLVRKQNELRLQGELODK 599
Db 620 NKIRDSYAQLLGHQNLKQKIKHVVKLDKNSQLKSEVSKLRSOLVRKQNELRLQGELODK 679
QY 600 ALGIRHFDPSKAFCHASKENFTPLKEGPNCC 630
Db 680 VLGIKHFDPKAFCHASKENFTPLKEGPNCC 713

RESULT 3
S67593
transport protein USO1 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein D2552; protein YDL058w
C:Species: Saccharomyces cerevisiae
C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 21-Jul-2000
C:Accession: S67593; A38455; S30782
R:Blöcker, H.; Brandt, P.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67597
A:Accession: S67593
A:Molecule type: DNA
A:Residues: 1-1790 <BLO>
A:Cross-references: EMBL:Z74106; NID:g1431058; PID:e253003; PID:g1431059; MIPS:YDL058
A:Experimental source: strain S288C
R:Nakajima, H.; Hirata, A.; Ogawa, Y.; Yonehara, T.; Yoda, K.; Yamasaki, M.
J. Cell Biol. 113, 245-260, 1991
A:Title: A cytoskeleton-related gene, USO1, is required for intracellular protein tra
A:Reference number: A38455; MUID:91185402; PMID:2010462
A:Accession: A38455
A:Molecule type: DNA
A:Residues: 1-389, 'TA', 392-724, 'S', 726-1790 <NAK>
A:Cross-references: GB:X54378; NID:g4777; PID:CAA38253.1; PID:g4778
A:Note: the authors translated the codon ACT for residue 768 as Ile
R:Hostetter, M.K.; Herman, D.J.; Bendel, C.M.; McClellan, M.; Tao, N.; Kendrick, K.E.
submitted to the EMBL Data Library, February 1993
A:Description: An integrin analogue in Saccharomyces cerevisiae.
A:Reference number: S30782
A:Accession: S30782
A:Molecule type: DNA
A:Residues: 71-846, 'E', 848-923, 'K', 925-1252, 'I', 1254-1318, 'V', 1320-1460, 'S', 1462-1580

A;Cross-references: EMBL:L03188
C;Genetics:
A;Gene: SGD:USO1; INT1
A;Cross-references: SGD:S0002216; MIPS:YDL058w
A;Map position: 4L
C;Keywords: coiled coil; transmembrane protein
F;326-342/Domain: transmembrane #status predicted <TM1>
F;394-410/Domain: transmembrane #status predicted <TM2>
F;617-633/Domain: transmembrane #status predicted <TM3>

Query Match 13.9%; Score 431.5; DB 2; Length 1790;
Best Local Similarity 24.2%; Pred. No. 5.5e-08;
Matches 179; Conservative 149; Mismatches 236; Indels 177; Gaps 33;

```
Qy 10 KLRNKRKTRKMRMMVKQEGMELKQATQKDLTESKGIQVLEGKLVSTIEKEKIDEKCE 69
Db 798 ELKNVRDS-LDEMQLRDLVLETKDKENQATLLEYKSTIHQEDSIKTEK----- 846
Qy 70 KLEYIQEISCASQVQCKVDIAQLEEDL-----KEKD-----REILSLKQ 111
Db 847 GLETILSQKKAEDGINKMGKDLFPALSRMQAVEENCKNQKEDKDSNVNHQKETSILKE 906
Qy 112 SLEENIT---FSQIEDLTVKCOLLETER-----DNLVSKDRERAET 150
Db 907 DIAAKITEIKAINLENLEEMKIQCNLNSKEKEHISKELVEYKSRFSQSHDNLVAKLITEKLS 966
Qy 151 LS---AEMQILTERL--ALERQEYE-KLQKELQSQ-SLLQKELKSLARLQOOLCSFOEE 203
Db 967 LANNYKDMQAEENSLIKAVEESKNESSTOLSNLKNKIDNSQKE-----NQIE 1016
Qy 204 MTS-EKNVFEELKALAEELDAVQOK-----ESESRLVQLEERKSTAPQL 250
Db 1017 RGSIEKNI--EQLKTTISDLQTEETIISKSDSKDEYESQISLLKEKLETATTANDENV 1074
Qy 251 TRDNLRLREKEVELEKHTAAHAQAILTAQEKYNDTAQSLRDVTAQLESVQV-----KY 303
Db 1075 NKISLTKTRF-ELEAELAAAYKLNKNELETKETSEKALKVEKNEHLKKEKIQLEKEA 1133
Qy 304 NDTAQSRLDVTQALSEQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRD-VTA---QL 359
Db 1134 TETQQLNSLRANLESLEKEHEDLAQULKYEEQIANKERYNEEISQINDIEITSTQOEN 1193
Qy 360 ESQVEKYND---TAQSLRDVSAQLESYKSTLK-----EIEDLKLLENLT----- 400
Db 1194 ESIRKKNDELEGEYKAMKSTSEESQNLKKSBDALNLQIKELKKKNETNEASLESISKV 1253
Qy 401 -----LQEKYMAEKSVEDVQOQILTAESTNQEYARMVQDLONRSTLKEEIKEIT 451
Db 1254 ESETVKIKELQDECNFKEVESELEDKLKASEDNKSKYL-----ELQKESEKIKBELDAKT 1309
Qy 452 SSF---LEKITDLKNLRQODEDFRQOLEEKGRKTAENVMTELTMINKWRLLYEELY 508
Db 1310 TELAIQLEKITNL-----SKAKESESE--LSRLKKTSSERKNNAEBQL 1351
Qy 509 EKTQ-PFQOQLDAFAEKQALLNEHGAT-----QEQLNKIRDSYAQLIGHONLKQIKHV 562
Db 1352 EKLNEIQIKNOAQEKERK-LLNGESSITQYSEKINTLDEILRLQENELKAK----- 1406
Qy 563 VKLDENSOLKSEYKSLRSQLVKQNELR-LQEL-----DKALGI-----RH 605
Db 1407 -EIDNTRSELE-KVSLNDELLEKQNTIKSLQLODEILSYKDKITRNDEKLLSIEDRNKR 1464
Qy 606 FDPKAFCHASKENFTPLKEG 626
Db 1465 LESLKEQLRAAQESKARVEEG 1485
```

RESULT 4
A59293
skeletal myosin heavy chain - domestic rabbit
C;Species: Oryctolagus cuniculus
C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 08-Sep-2000
C;Accession: A59293

R;Maeda, K.; Hostinova, E.; Roesc.Klein Kauf, A.; Schuster, H.; Gasperik, J.; Wittingh
submitted to GenBank, July 1995
A;Description: Isolation, sequencing of myosin heavy chain cDNA from rabbit skeletal
A;Reference number: A59293
A;Accession: A59293
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1938 <MAE>
A;Cross-references: GB:U32574; NID:g940232; PIDN:AAA74199.1; PID:g940233
A;Experimental source: strain New Zealand White; cell type skeletal muscle fiber type
C;Genetics:
A;Gene: MHC
C;Superfamily: myosin heavy chain; myosin motor domain homology
F;89-769/Domain: myosin motor domain homology <MCO>

Query Match 13.6%; Score 421; DB 2; Length 1938;
Best Local Similarity 23.0%; Pred. No. 1.4e-07;
Matches 170; Conservative 150; Mismatches 251; Indels 168; Gaps 25;

```
Qy 9 MCLRNRKTRKMRMMVKQEGMELKQATQKDLTESKGIQVLEGKLVSTIEKEKIDEK 66
Db 833 MKLYFKIKPLLKSAETEKEMANMKKEFEKTLAKAKAEKAEKELKEMVALMQEKNDLQL 892
Qy 67 ETEKLEYIQEISCASQVQCKV-----DIAQLEEDLKEKREILSLKOSLEENIT- 118
Db 893 QVQAEADSLADAEERCQDLIKTIQLEAKIKIEVTERAEDEEEINAEITAKKRKLEDECE 952
Qy 119 FSKQIEDLTVKCOLLETERDNLVSKDRERAETLSAEMOILTERLALEQVEYKELQOKELO 178
Db 953 LKDDIDDLUETLAKVEKEKHATEN---VKNLTTEMAGLDETIAKLTKEKALQEAHQO 1008
Qy 179 QSLSLQKEKE-----LSARLQOOLCSFOEEMTSEKNV-----FKBELKIAL-A 220
Db 1009 TLDDQLAEEDKVNTLTAKTKLEQQVDDLEGSLEQEKIRMDLERAKRKLEGDLKLAQES 1068
Qy 221 ELDAVQKKEEESERLVKQ-----LE---EERKSTAEQILTRLDNLRLKEVELEKHIAAH 271
Db 1069 TMDIENDRQQQIDELKKKEFEFENLSQKIEDQALAMOLQKKIKELQARIEELEEEIEAE 1128
Qy 272 AQAILTAQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESQEKYNDTAQSL 331
Db 1129 RASRAKAEKQSDJUSLEEEISERLE-----EAGGATSAQIEMNKKREAEFQKMR 1178
Qy 332 RDVTAQLESQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESYKSTLKREI 391
Db 1179 RD---LEEATLQHEATAATLRKHA--DSVAE-LGEQIDNLRQVKQLEKEKSELKWEI 1231
Qy 392 EDLKLLENLTQEKYMAEKSVEDVQOQILTAESTNQEYARMVQDLO-NRSTLKEE----- 445
Db 1232 DLLASNMTVSKAKGNLEKMCRTLEDQVSELKTBEEHQRLINDLSAQRARLQTESGEFS 1291
Qy 446 -----EIKETSSFEKLTIDLKNLRQO-----DED-FRQOLE 477
Db 1292 ROLDEKDSLVSGKQGAFTQQIEELKRQLEEEITAKSALAHALQSAHRHCDLLREQYE 1351
Qy 478 EKGKRTAEKENVMTELTMEINKWRLLY-----EELYETKPFQOQLD----- 519
Db 1352 EEQAKAELQRAMSKANSEVAQWRKYETDALQRTVEELEEAKKLAQRLODAEEHVAVN 1411
Qy 520 ----AFAEKQALLNE-----HGATQEQNLN-----KTRDSYAQLL 550
Db 1412 AKCASLEKTKORLQNEVEDLMIDVERTNACAALDKQKQNFQKILAEWKHKYEETHAELE 1471
Qy 551 GHONLKQ-----KIKHVVK-----LKDENSOLKSEYKSLRSQ----- 583
Db 1472 ASQKESRSLSTEVFKVKNAYEESLDQLETLKRENKNLQOETSIDLTEQIABEGKRIHELEK 1531
Qy 584 VKR--KQNELRLQGEELDKA 600
Db 1532 VKQVEQEKSELQAALEEA 1550
```

RESULT 5

A57013
n:Early endosome antigen 1 - human
C:Alternate names: endosome-associated protein
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A57013; S44243
J. Biol. Chem. 270, 13503-13511, 1995
R. Mu, F.T.; Callaghan, J.M.; Steele-Mortimer, O.; Stenmark, H.; Parton, R.G.; Campbell, R.; Biol. Chem. 270, 13503-13511, 1995
A:Title: EEAL1, an early endosome-associated protein. EEAL1 is a conserved alpha-helical p
A:Reference number: A57013; MUID:95286647; PMID:7768953
A:Accession: A57013
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1410 <RES>
A:Cross-references: GB:L40157; NID:g1016367; PIDN:AAA79121.1; PID:g1016368
R:Seelig, H.P.
submitted to the EMBL Data Library, April 1994
A:Reference number: S44243
A:Accession: S44243
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-254, 'C', 256-257, 'LQ', 260-276, 'A', 278-283, 'A', 285-519, 'D', 521-574, 'EQ', 577-
A:Cross-references: EMBL:X78998; NID:9475933; PIDN:CAA55632.1; PID:9475934
C:Genetics:
A:Gene: GDB:EEAL
A:Cross-references: GDB:1369996
C:Superfamily: human early endosome antigen 1
C:Keywords: calmodulin binding; endocytosis; metal binding; peripheral membrane protein;

Query Match 13.4%; Score 415; DB 1; Length 1410;
Best Local Similarity 22.8%; Pred. No. 1.6e-07;
Matches 168; Conservative 153; Mismatches 223; Indels 192; Gaps 28;

Qy 1 MRALSLEMLKLRN-----KRETKRSMVMVKQEG---WELK-----LQAT--OK 38
Db 263 VRVYVQELQKLSVSSNETQKNTUTENLLKKEQDYTKLEKHNEESVSKNIIQATLHQK 342
Qy 39 DITSGKIVQLGKLVSIKE-----KIDKCB-TEKLEYIOEISCASDQVEKCKVD 91
Db 343 DL-----DCQQLQSLASSETSLSRIHVSEKGEATQKLEELSEV---ETKYQHLKAE 394
Qy 92 IAQLEDLKDREILTSKLSQLEENITTSKQIEDITVVCQQLTERDNLVSKDRERATL 151
Db 395 FKLOQQQREKEQHGLQSEINQ-----LHKKLETERQLGEAHR-----L 437
Qy 152 SAEMQILTERLALEREQYEKQKELQSQSLQOEKLSARLQOL-----CSFQ 201
Db 438 KEORLSSEKLMDEQVADQLKLSRLEEQLKEKVTNSTLOHQDLTKQHQEQQAQ 497
Qy 202 EEMTSEKNVFEELKALAEILDVQOEQSERLVKQ-----LEERK----- 244
Db 498 QSTTAKLREAQNDLEQVLRQIGEKDQKQNTLEALLQSKENISLLEKEREKEDLYAKIQAGE 557
Qy 245 ---STAEQLTRLDNLLREKEVELEKHIAHAQAIIIAQEKYNDTAQSLRDVTAQLESVQE 301
Db 558 GETAVNLQEKNHFTLQDEVTLQNTENVNKNQSESHKQAQENLHDQVQEQK---AHLRAAQD 614
Qy 302 KYNDTAQSLRDVTAQLESQEKYND------TAQSLRDVTAQLESQ 342
Db 615 RVLSLTSVNLNSQLNESKESKVSQSLDIQAKTELLLSABAATAQR-ADLNHLDTAQ 673
Qy 343 EKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVSAQLESYKSSTL-----KEIDL----- 394
Db 674 NALQDKHQELNKIITQLQDVVTAQKLQDKQEHCSQLESHLKEYKELYSLQEKTEELGOIK 733
Qy 395 KLENTLTQEKVAMARKSVEDVQQQIILTAESTNQEVARMVQDLQNRSTLKEBEI---KEIT 451
Db 734 KLEADSEV-KASKEQALQDLQEQ-----RQLNTDLERATLSKQLEMEKEIV 781
Qy 452 SSF-----LEKTDKLNQLRQOQEDFRKOLEEGKKTAEKNVMTLTWTWINKRWLLY 504
Db 782 SSTRLDLQKKSEALBSIKQKTKQEE--KOILKQDFTLSQET-----KIQH 827

Qy 505 EELYEKTRPFOQQLDAFAEKQKALLNEHGATQEQINKIRDSY-----AQL 549
Db 828 EELNNRIQTIVTQKLVKMEKALMTLSTVTKDKLSKVSLSKNSKFEKENQKGA 887
Qy 550 LGHONLQKIKHVVVKLDENS-----QLKSEVSKLRSOLVVR----- 586
Db 888 LDLEKTKELCHQLQVOMENTLKQEKLSLEKEKASHQKLEKLSMQEQLIQAOQNTL 947
Qy 587 KON---ELRLQGELEDK 599
Db 948 KQNEKEEQLOQNGINE 963

RESULT 6
A24922
myosin heavy chain, skeletal muscle, embryonic - rat
N:Contains: myosin ATPase (EC 3.6.4.1)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 19-Apr-2002
C:Accession: A24922; A22538; B24263
R:Strehler, E.E.; Strehler-Page, M.A.; Perriard, J.C.; Perlasamy, M.; Nadal-Ginard, B
J. Mol. Biol. 190, 291-317, 1986
A:Title: Complete nucleotide and encoded amino acid sequence of a mammalian myosin he
A:Reference number: A24922; MUID:87060988; PMID:3783701
A:Accession: A24922
A:Molecule type: DNA
A:Residues: 1-1940 <STR>
A:Cross-references: GB:X04267; GB:X05004; NID:g56658; PIDN:CAA27817.1; PID:g1619338
R:Strehler, E.E.; Mahdavi, V.; Perlasamy, M.; Nadal-Ginard, B.
J. Biol. Chem. 260, 468-471, 1985
A:Title: Intron positions are conserved in the 5' end region of myosin heavy-chain ge
A:Reference number: A22538; MUID:85080119; PMID:2981212
A:Accession: A22538
A:Molecule type: DNA
A:Residues: 1-168 <STR>
A:Cross-references: GB:L00370; GB:M0135; NID:g205580; PIDN:AAA41655.1; PID:g9554476
R:Perlasamy, M.; Wydro, R.M.; Strehler-Page, M.A.; Strehler, E.E.; Nadal-Ginard, B.
J. Biol. Chem. 260, 15856-15862, 1985
A:Title: Characterization of cDNA and genomic sequences corresponding to an embryonic
A:Reference number: A24263; MUID:86059474; PMID:2999140
A:Accession: B24263
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1358-1490, 'G' <PER>
A:Cross-references: GB:K03468; NID:g205573; PIDN:AAA41652.1; PID:g205574
A:Experimental source: clone pMHC-72
C:Genetics:
A:Introns: 68/3; 116/3; 169/1
A:Note: the list of intron positions may be incomplete
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; muscle
F:89-767/Domain: myosin motor domain homology <MMOT>
F:179-186/Region: nucleotide-binding motif A (P-loop)
F:549-586/Region: actin binding #status predicted
F:656-678/Region: actin binding #status predicted
F:840-1940/Domain: coiled coil #status predicted <COI>
F:840-1280/Region: S2
F:1281-1940/Region: light meromyosin
F:130/Modified site: N6,N6-trimethyllysine (Lys) #status predicted
F:185/Binding site: ATP (Lys) #status predicted
F:696,706/Active site: Cys #status predicted

Query Match 13.4%; Score 415; DB 1; Length 1940;
Best Local Similarity 22.0%; Pred. No. 2.2e-07;
Matches 165; Conservative 144; Mismatches 244; Indels 198; Gaps 22;

Qy 4 LSLEMLKLRNKRRTKMRSMVMVKQEGMELKLAQTKDLTESKGIIVQLEGKLVSTKEKID 63
Db 833 LFFKIKPLKLSAETKEKEMATMKEE-----FQTKDELAKSEAKRKELEKLVTLVQEKND 887
Qy 64 ----EKCEKLELYIQEISCASDQVEKCV-----DIAQLEDLKEKDEILSLKQS 112
Db 888 LQLOVQAENSELLID--AEERC--DQLIKAKFQLEAKIKVEVTERAFDEDEEINAEITAKRRK 943

Qy 113 LEENIT-FSKQIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQIILTERLALRQEQYEK 171
| : : | : | : : | : | : : | : | : : | : | : : | : | : : | : | : :
Db 944 LEDECSSELKRDIDDLTLAKVEKEKATENK-----VKNLTELGLAGDEYIAKLTREKKA 999
| : : | : | : : | : | : : | : | : : | : | : : | : | : : | : | : :
Qy 172 LQKELQSLSLLOEKE-----LSARLQOQLSFOEMTSEKNV-----FKEE 214
| : : | : | : : | : | : : | : | : : | : | : : | : | : : | : | : :
Db 1000 LQEAHQQTLDLQAEEDKVNLSKLKLEQOQVDLSESSLEQEKKLAVDLERNKRKLEGD 1059
| : : | : | : : | : | : : | : | : : | : | : : | : | : : | : | : :
Qy 215 LKALAE-LDAVQOKEQSRELVK-----OLE-----ERKSTAEQLTRDNLNLLREKEVEL 264
| : : | : | : : | : | : : | : | : : | : | : : | : | : : | : | : :
Db 1060 LKLAQESILDENDQDLDEKLKKDFEYSOLOSKVDEQTLQLOQKKTKELQARTEEL 1119
| : : | : | : : | : | : : | : | : : | : | : : | : | : : | : | : :
Qy 265 EKHTAAHAQALLIAQKYNNTAQSLRDVTAQLESVQEKYNNTAQSLRDVTAQLESEQ-EK 323
| : : | : | : : | : | : : | : | : : | : | : : | : | : : | : | : :
Db 1120 EEEIEAERATRAKTEKQDSYARELEELSERLEE-----AGGVSTQIELNKKREAE 1171
| : : | : | : : | : | : : | : | : : | : | : : | : | : : | : | : :
Qy 324 YNDTAQSLRDVTAQLESE-----QEKYNNTAQSLRDVTAQLESVQEKYNNTAQSLRDVSAQ 379
| : : | : | : : | : | : : | : | : : | : | : : | : | : : | : | : :
Db 1172 FLKLRDLEAATLQHEATVATLRKKHADSAELAEQIDNLRQVKOK----- 1217
| : : | : | : : | : | : : | : | : : | : | : : | : | : : | : | : :
Qy 380 LESYKSTLKEIEDKLLENLTLOEKVMAEKSVEDVQOQIILTAESTNQEVYARVQDLONR 439
| : : | : | : : | : | : : | : | : : | : | : : | : | : : | : | : :
Db 1218 LEKESKFKLEIDLLSSVESVSKANLEKICRTLEDQLESEARGKNEETQRSLSELTTQ 1277
| : : | : | : : | : | : : | : | : : | : | : : | : | : : | : | : :
Qy 440 STLKEEIKETIT-----SSFLEKITDLKNLRQOQED----- 471
| : : | : | : : | : | : : | : | : : | : | : : | : | : : | : | : :
Db 1278 KSRLOTAGELSRQLEBKESIVSLSRSKQAFQTOIEELKQLEENKAKNALAHALQSS 1337
| : : | : | : : | : | : : | : | : : | : | : : | : | : : | : | : :
Qy 472 -----FRKQLEEKGRKTAENYMTLMTBEINKWRLLY-----PELYEKTTPFQQQ 517
| : : | : | : : | : | : : | : | : : | : | : : | : | : : | : | : :
Db 1338 RHDCDLLREQVEEQEKAELQALSKANSEVAQWRTKYETDAIQRTTEELBEAKKKLAQR 1397
| : : | : | : : | : | : : | : | : : | : | : : | : | : : | : | : :
Qy 518 L----- 518
| : : | : | : : | : | : : | : | : : | : | : : | : | : : | : | : :
Db 1398 LODSEEQVEAVNAKCALEKTKQRLQGEVEDLMVDVERANSAAALDKQKORNFQKVLAEW 1457
| : : | : | : : | : | : : | : | : : | : | : : | : | : : | : | : :
Qy 519 ----DAFEAEKQALLNHEGATQEQOLNKRIRDSYAQLLGHQNLKQIKKHVKLKDNSOLKS 574
| : : | : | : : | : | : : | : | : : | : | : : | : | : : | : | : :
Db 1458 KTKCEEQAELEALKESRSLSTELFKLNAY-----EELDQLQETVKRENKLEQ 1508
| : : | : | : : | : | : : | : | : : | : | : : | : | : : | : | : :
Qy 575 EVSKLSQLVKRNQK-----ELRLQOGLDKA 600
| : : | : | : : | : | : : | : | : : | : | : : | : | : : | : | : :
Db 1509 EIADLTEQIAENGKSIHELEKSRQKMELEKA 1539
| : : | : | : : | : | : : | : | : : | : | : : | : | : : | : | : :

RESULT 7
JX0178
myosin heavy chain, fast skeletal muscle, adult [validated] - chicken
N:Contains: myosin Appase (PC 3.6.4.1)
C:Species: Gallus gallus (chicken)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 19-Apr-2002
C:Accession: PX0050; PX0051; PX0052; JX0178; A26365; S02082; PW0009; S39081; S24351; S05515
R:Hayashida, M.; Maita, T.; Matsuda, G.
J. Biochem. 110, 54-59, 1991
A:Title: The primary structure of skeletal muscle myosin heavy chain: I. Sequence of the
A:Reference number: PX0050; MUID:92041767; PMID:1939027
A:Accession: PX0050
A:Molecule type: protein
A:Residues: 1-205 <HAY>
R:Komine, Y.; Maita, T.; Matsuda, G.
J. Biochem. 110, 60-67, 1991
A:Title: The primary structure of skeletal muscle myosin heavy chain: II. Sequence of th
A:Reference number: PX0051; MUID:92041768; PMID:1939028
A:Accession: PX0051
A:Molecule type: protein
A:Residues: 206-636 <KOM>
R:Maita, T.; Miyanishi, T.; Matsuzono, K.; Tanioka, Y.; Matsuda, G.
J. Biochem. 110, 68-74, 1991
A:Title: The primary structure of skeletal muscle myosin heavy chain: III. Sequence of t
A:Reference number: PX0052; MUID:92041769; PMID:1939029
A:Accession: PX0052
A:Molecule type: protein

A:Residues: 201-213;632-837 <MAI>
R:Maita, T.; Yajima, E.; Nagata, S.; Miyanishi, T.; Nakayama, S.; Matsuda, G.
J. Biochem. 110, 75-87, 1991
A:Title: The primary structure of skeletal muscle myosin heavy chain: IV. Sequence of
A:Reference number: JX0178; MUID:92041770; PMID:1939030
A:Accession: JX0178
A:Molecule type: protein
A:Residues: 833-1938 <MA2>
R:Maita, T.; Hayashida, M.; Tanioka, Y.; Komine, Y.; Matsuda, G.
Proc. Natl. Acad. Sci. U.S.A. 84, 416-420, 1987
A:Title: The primary structure of the myosin head.
A:Reference number: A26365; MUID:87092420; PMID:3467365
A:Accession: A26365
A:Molecule type: protein
A:Residues: 1-129, 'X', 131-139, 141-550, 'X', 552-754, 'X', 756-784, 'QL', 787-804, 806-810 <M
R:Watanabe, B.
Biochem. Chem. Hoppe-Seyler 370, 55-61, 1989
A:Title: Amino-acid sequence of the hinge region in chicken myosin subfragment-2.
A:Reference number: S02082; MUID:89228549; PMID:2713098
A:Accession: S02082
A:Molecule type: protein
A:Residues: 1144-1270 <WAT>
R:Yajima, E.
Nagasaki Igakkai Zasshi 65, 409-430, 1990
A:Title: Study on tail region of skeletal muscle myosin; primary structure and protea
A:Reference number: PW0009
A:Accession: PW0009
A:Molecule type: protein
A:Residues: 1304-1938 <YAJ>
R:Moore, L.A.; Arrizubieta, M.J.; Tidyman, W.E.; Herman, L.A.; Bandman, E.
submitted to the EMBL Data Library, August 1991
A:Description: Analysis of the chicken fast myosin heavy chain family: Localization o
A:Reference number: S39081
A:Accession: S39081
A:Molecule type: mRNA
A:Residues: 1081-1203, 'DV', 1206-1342, 'E', 1344-1544, 'S', 1546-1795, 'HV', 1798-1829, 'S', 1
A:Cross-references: EMBL:M74084
R:Moore, L.A.; Arrizubieta, M.J.; Tidyman, W.E.; Herman, L.A.; Bandman, E.
J. Mol. Biol. 225, 1143-1151, 1992
A:Title: Analysis of the chicken fast myosin heavy chain family. Localization of isof
A:Reference number: S24348; MUID:92309413; PMID:1377278
A:Accession: S24351
A:Molecule type: mRNA
A:Residues: 1082-1182, 'T', 1184-1203, 'DV', 1206-1342, 'E', 1344-1544, 'S', 1546-1780, 'M', 17
A:Cross-references: EMBL:M74084
R:Watanabe, B.
Biochem. Chem. Hoppe-Seyler 370, 1027-1034, 1989
A:Title: Complete amino-acid sequence of subfragment-2 in adult chicken skeletal musc
A:Reference number: S05515; MUID:90121764; PMID:2610940
A:Accession: S05515
A:Molecule type: protein
A:Residues: 842-906, 'Q', 908-1270 <WA3>
R:Watanabe, B.
Biochem. Chem. Hoppe-Seyler 370, 549-558, 1989
A:Title: Amino-acid sequence of the short subfragment-2 in adult chicken skeletal mus
A:Reference number: S04501; MUID:89374803; PMID:2775482
A:Accession: S04501
A:Molecule type: protein
A:Residues: 852-906, 'Q', 908-1108 <WA2>
R:Matsuda, G.; Maita, T.; Miyanishi, T.; Hayashida, M.
J. Protein Chem. 6, 33-46, 1987
A:Title: Structure and function of muscle myosin.
A:Reference number: A60877
A:Accession: A60877
A:Molecule type: protein
A:Residues: 1-139, 141-205 <NAL>
R:Gulick, J.; Kropp, K.; Robbins, J.
J. Biol. Chem. 260, 14513-14520, 1985
A:Title: The structure of two fast-white myosin heavy chain promoters. A comparative
A:Reference number: A92507; MUID:86033956; PMID:2997212
A:Accession: A24124
A:Molecule type: DNA
A:Residues: 'M', 1-168 <GUL>

A:Cross-references: GB:M13512; GB:M12083; GB:M13510; NID:G212363; PIDN:AAA48966.1; PID:G
E:Kropp, K.; Gulick, J.; Robbins, J.
J. Biol. Chem. 261, 6613-6618, 1986
A:Title: A canonical sequence organization at the 5'-end of the myosin heavy chain genes
A:Reference number: A92587; MUID:86196091; PMID:3009465
A:Accession: C25217
A:Molecule type: DNA
A:Residues: 'M',1-56, 'T',58-76, 'I',78-168 <KRO>
A:Cross-references: GB:M13515; GB:M13511; NID:G212373; PIDN:AAA48971.1; PID:G553468
C:Comment: This is a fragment of the globular head.
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: acetylated amino end; actin binding; ATP; coiled coil; hydrolase; methylated
E:1-1938/Product: myosin heavy chain #status experimental <MAT>
F:89-768/Domain: myosin motor domain homology <MMOF>
F:179-186/Region: nucleotide-binding motif A (P-loop)
F:550-587/Region: actin binding #status predicted
F:657-679/Region: actin binding #status predicted
F:841-1938/Domain: coiled coil <COI>
F:841-1289/Region: S2
F:852-1108/Domain: short subfragment 2 <SUB2>
F:1290-1938/Region: light meromyosin
F:1/Modified site: acetylated amino end (Ala) #status experimental
F:35/Modified site: N6-methyllysine (Lys) #status experimental
F:130-551/Modified site: N6,N6,N6-trimethyllysine (Lys) #status experimental
F:185/Binding site: ATP (Lys) #status predicted
F:697,707/Active site: Cys #status predicted
F:755/Modified site: 3'-methylhistidine (His) #status experimental

Query Match	13.2%	Score 411;	DB 1;	Length 1938;
Best Local Similarity	22.6%	Pred. No. 3e-07;		
Matches 171;	Conservative 135;	Mismatches 247;	Indels 204;	Gaps 25

Qy 9 MCLRNRKTRMSVMVQ--GMEJLQATQKDLTSGKIVQLEGKIVSIEKEIKDEKC 66
 ||| } : : : || : : : || : : : || : : : || : : : || : : : || : : : || : : :
 Db 832 MKLFFKIPLLKSAESEKEMANMKFEFEKTEELAKSEAKRKELEEMVVLQEKNDIQL 891

Qy 67 ETEKLEYIQEISCASDQVEKCV-----DIAQEEDLAEKDREITLSLKQSLEENIT- 118
Db 892 QVQAADSLDAERCDQLIKTKIQLEAKTKEVTERADEEEINAEUTAKKRLEDCESE 951

Qy 119 FSKQIEDLTVKCOLLETEDNRNLVSKDRERAFTLSAEMQILTLRLALEREQEYKQLQKEIQ 178
 | : || : | : : | : || : || : | : | : | : | : | : | :
Db 952 LKKDIDDLTLLAKEVEKEKHATENK-----VKNLTEAMAVLDETIAKTKEKKAQEAHQ 1007

Qy	179	QSULLQEQKE-----LSARLQQQLCSFQEMTSEKNV-----	210
		: : : : : : : :	
Db	1008	TLDLQVEEDKVNLTAKTKLEQVDDLEGSLEQEKKLRMDLERAKRKLEGLDKLANDS	1067

QY 211 -----FKEELKLAELDVAQQEEOSERLVKKOLEERKSTAEQLTRLDNLLBK 260
 :
Db 1068 IMDLENDKQOLDEKLKKKFISIQSKIEDOALGMOLQRKIK ---ELOARTEEL---EE 1122

QY 261 EVELEKHTAAHAQAIIIAOEKYNDTAQSLRDVTAQLESVQEKNYDNTAQSLRDVTAQLESE 320
 :
Db 1123 EIEAERTSRAKA-----EKHR-----ADLSRELEIISELEEAGCA---TAAOIEMN 1166

```
QY 321 QEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLR-----DVTQAQLESVOEKYNDTAQSLRDV 376
      :: : : || || :: || :: || || :: || :: || :: || :: || :: || :: || :: ||
Db 1167 KKREAEFOKMRD-----LEATLQHEATAAALRKKHADSTAFI-----GEQIDNLRV 1215
```

Qy 377 SAQLESYSSSTLKEIEDLKLENTLQEKVAAEK---SVEDVQQQLTAESTNQEYARMV 433

Db 1216 KQLEKEKSELKWEITDDIASNMSYSYSAKANEKMCRTIPEQISETKTKFEONO---BMT 1272

Qy	434	QDLQRRSTLKEEETKEIT-----SSFLEKITDLKNLQRRQ-----	468
Db	1273	NDINTORARIOTGTGEYSROAEFKKDALISQSRGKGQETQTFEELKSHLEETKAKNAIA	1332

Qy 469 -----DED-PRKQLEKGKRTAEKENVNTMTLMEINKWRLLY-----EELYEKT 511
| | : | : | : | : : : | : |||
Db 1333 HALOSARHDCDLIRFOYVEFFOFAKGELORAI-SKANSEVAOMETKYTFDAJORTFEELERAK 1302

Qy 512 KBFQQQLD-----AFEAKEQALLNE-----HGATQEQLNKIRDSYA 547

A:Cross-references: GB:M13512; GB:M12083; GB:M13510; NID:g212363; PIDN:AAA48966.1; PID:g212363; GB:M13515; GB:M13511; NID:g212373; PIDN:AAA48971.1; PID:g555468
 R:Kropp, K.; Gullick, J.; Robbins, J.
 J. Biol. Chem. 261, 6613-6618, 1986
 A:Title: A canonical sequence organization at the 5'-end of the myosin heavy chain genes
 A:Reference number: A92587; MUID:86196091; PMID:3009465
 A:Accession: C25217
 A:Molecule type: DNA
 A:Residues: 'M', 56, 'T', 58-76, 'I', 78-168 <KRO>
 A:Cross-references: GB:M13515; GB:M13511; NID:g212373; PIDN:AAA48971.1; PID:g555468
 C:Comment: This is a fragment of the globular head.
 C:Superfamily: myosin heavy chain; myosin motor domain homology
 C:Keywords: acetylated amino end; actin binding; ATP; coiled coil; hydrolase; methylated
 F:1-1938/Product: myosin heavy chain #status experimental <MAT>
 F:89-768/Domains: myosin motor domain homology <MMOT>
 F:179-186/Region: nucleotide-binding motif A (P-loop)
 F:550-587/Region: actin binding #status predicted
 F:657-679/Region: actin binding #status predicted
 F:841-1938/Domains: coiled coil <COI>
 F:841-1289/Region: S2
 F:852-1108/Domains: short subfragment 2 <SUB2>
 F:1290-1938/Region: light meromyosin
 E:1/Modified site: acetylated amino end (Ala) #status experimental
 F:35/Modified site: N6-methyllysine (Lys) #status experimental
 F:130,551/Modified site: N6,N6,N6-trimethyllysine (Lys) #status experimental
 F:185/Binding site: ATP (Lys) #status predicted
 F:697,707/Active site: Cys #status predicted
 F:755/Modified site: 3'-methylhistidine (His) #status experimental

Query Match 13.2%; Score 408.5; DB 2; Length 1690;

Best Local Similarity 22.2%, P-Seq. NO. 3.2e-07,
Matches 174; Conservative 162; Mismatches 240; Indels 209; Gaps 31;

Qy	1	MRALSLELMKLRN-KRET	KRSMRMVMVQEGH	NELKQATQKDLT	ESKGVILEGKLV----	55
D6	543	LEKTSLSGECIENLRRL	ALLKEENKQAEQA	AEFTRK-LAEKSV	EVLSLSLSEQLNKA	601

Qy	56	---	SIEKEKID-EKCE-----	TEKLLEYTQEISASDOVECKVKVIAQLBEDLK-	100
Db	602	:	: : : :	TSDLESERYNKSDBCEIILQTEVRMRDEQIRELNQOLDEVYTLNQVKADSSALDMDMLRL	661

Qy	101	-----EKDREILSLQ-----SLEENITFSKQIEDLTAVKCOLLETERDNLVSK	143
Db	662	QKEGTEEKSTLLEKTEKELVQIKEQAATKLDQKEQLEKQISDLK---QLAQEKEK	716

```

144 DRERA-ETLSAEMQILTERALEREQEYKLOQKELQSLSLQOEK-----E 188
      |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
717 KTEAINOIOLEKESIEOOLAKONELEDFOKQOSESEVHLEIKAQNTOKDLEVSQGE 776
      |   |   |   |   |   |   |   |   |   |   |   |   |   |   |

```

```

QY 189  LSARLQQQLCSFQEEMTSEKNVFXEELKALAEALDAVQQKEEQSERLVKOLEEERKSTAE 248
      :|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 777  SLKKLOOL-----EKEITGHEKLOAAEEL-----KKETIIRKEOELOOLOOS 822
      :|||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

QY 249 QLRLDNLRLREKEVELEKHIAHAQAAILAQEKYNDTAQLRDVTAQL-----ESVQEKYN 304

Db 823 KSAESESALKVVOLEO-----LOOAAAASGEGKTVAKLHDET SOLKSOAEFTOSELK 878

QY	305	DTAQSRLDVTAQLSESEQEKEYNDTAQ---	SLRDTVTAQLSESEQE-----	YNDTAQS----	351
Dh	879	STENLEAKSKOLRAANGSIFEEAKKSGOLOPITTKIKSEVEFTAAALSVVHTDVFESKTK			938

QY	352	-LRDVTQAQLSVQEKYNDT--AQLSRD-----VSAQLSYKSSLT-----	388
Dh	939	CLFAANAALAEKVNVEYAEAFSAASLQDVKETDTTLHAFLOAPRSSSSSAHTKISKESD	998

```

QY 389 -----KEIDLKLENTLQ-----KVMAAEKSVE 413
          : : : : :
999 FIATCHKETSKADAWSOEMOKFELOLPOOLODSOSOTKIKAEGEKEFEESIK 1058
          : : : : :
Dh

```

QY 414 DVQQIILTAESTNOEYA----RMVQDLQNRSTLKEEI---KEITSSFLEKITDLKN--- 463

Db 895 QVOAEDAGLADAEBCDQIKTKIQLEAKIKELTERAEDEEENNAELTAKKKRLEDECS 954
QY 119 FSKQIEDLVKCOLLETERDNLVSKDRERAETLSAEMOILTERLALEREYKELQKELQ 178
Db 955 LKKDIDDLLETLAKVEKHEATENK----VKNLTEMAALDETIAKTREKALQEAHQ 1010
QY 179 S-----QSLLOQKEL-----SARLOQOLCSFQB- 202
Db 1011 TLDDLQAEEDKVNLTAKTKLEQQVDDLEGSLEQEKRLMDLERAKKRLEGDLKMTQES 1070
QY 203 --EMTSEKNVFEELKLAELADVAQKEEQSERLVKQLEEEERKSTAEQTLRLDNLRLRK 260
Db 1071 TMDLENDKQOLDEKLKKDFEISQIQSKIEDEQALGMLOKKIK----ELQARIEEL--EE 1125
QY 261 EVELKHHAAQAAILIAOEKYN-DFAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLES 319
Db 1126 EIAEERTSRAK-----EKHRADLSRELEISERLE---TAGGATAAQI-DMKKRRA 1174
QY 320 EQKYNPTAQSLRDVTAQLESE-----QEKYNPTAQSLRDVTAQLESVQEKYNDTAQSLRD 375
Db 1175 EFQKMR---RDLEEATLQHEATAAALRKHK---ADSTADVGEQIDNLQR----- 1217
QY 376 VSAQLESYSSTLKETEDLKENLTLOEKVMAEK---SVEDVQOQILTAESTNQ----- 427
Db 1218 VRQKLEKSEKSELMEIDDLASNMESVSKAKANLEKMCRLSDQLSEIKTKKEEQOQTIND 1277
QY 428 -----EYARMVQDLQNRSTLKEEIKETT---SSFLEKITDLKNOLRQO--- 468
Db 1278 ISAKARLOTESGEYSRVQEE-----KDALISLSRGKQAFQTOIEELKRLHLEEIK 1330
QY 469 -----DED-FRKQLEEKGRTPAKENVMTEMTINKWRLLY-----EE 506
Db 1331 KKCPAHALQASRHDCDLLREQVEEQEAKGELQALSKANSEVAQWRTKYETDAIQRTTE 1390
QY 507 LYEKTRPFQOOLD-----APEAEKQALLNE-----HGATQOLNKI 542
Db 1391 LBEAKKKLAQRLQDAEBHVEAVNSKQASLEKTKQRLQNEVEDLMDIVERSNAACAALDKK 1450
QY 543 RDSYAQLLGHQNLK-----VRR--KQNELRLQGLDKA 600
Db 1451 QKNFDKILSEWKQYKBEETOAELEASQESRSLSLSTELFKMKNAYESLDHLETLKRENKL 1510
QY 573 KEVSKLRSQL-----VRR--KQNELRLQGLDKA 600
Db 1511 QOEISDLTEQIAEGGKAIHELEKVKKQIBQEKSELQTALEEA 1552

RESULT 11
S04090
myosin heavy chain 3, skeletal muscle, embryonic - human
N:Contains: myosin ATPase (EC 3.6.4.1)
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text change 19-Apr-2002
C:Accession: S04090; S06146; S05442; S12460; S09333; A35082
R:Eller, M.; Stedman, H.H.; Sylvester, J.E.; Fertels, S.H.; Rubinstein, N.A.; Kelly, A.M.
Nucleic Acids Res. 17, 3591-3592, 1989
A:Title: Nucleotide sequence of full length human embryonic myosin heavy chain cDNA.
A:Reference number: S04090; MUID:89263803; PMID:2726495
A:Accession: S04090
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-1940 <ELL>
A:Cross-references: EMBL:X13988; NID:g34843; PIDN:CAA32167.1; PID:g34844
R:Eller, M.; Stedman, H.H.; Sylvester, J.E.; Fertels, S.H.; Wu, Q.L.; Raychowdhury, M.K.
FEBS Lett. 256, 21-28, 1989
A:Title: Human embryonic myosin heavy chain cDNA. Interspecies sequence conservation of
A:Reference number: S06146; MUID:90033298; PMID:2806546
A:Accession: S06146
A:Molecule type: mRNA
A:Residues: 774-1662, 'OT', 1665-1940 <LEL>
A:Cross-references: EMBL:X13100; NID:g31143; PIDN:CAA31492.1; PID:g31144
R:Karsch-Mizrachi, I.; Travls, M.; Blau, H.; Leinwand, L.A.
Nucleic Acids Res. 17, 6167-6179, 1989

A:Title: Expression and DNA sequence analysis of a human embryonic skeletal muscle my
A:Reference number: S05442; MUID:89366648; PMID:2771643
A:Accession: S05442
A:Molecule type: DNA
A:Residues: 856-1390, 'KK', 1393-1940 <KAR>
A:Cross-references: EMBL:X15696; NID:g36504; PIDN:CAA33731.1; PID:g1335313
R:Stedman, H.H.; Eller, M.; Jullian, E.H.; Fertels, S.H.; Sarkar, S.; Sylvester, J.E.
J. Biol. Chem. 265, 3568-3576, 1990
A:Title: The human embryonic myosin heavy chain. Complete primary structure reveals e
A:Reference number: A35082; MUID:90154023; PMID:2303463
A:Contents: annotation; chromosomal assignment
R:Bober, E.
submitted to the EMBL Data Library, January 1989
A:Reference number: S12458
A:Accession: S12460
A:Molecule type: mRNA
A:Residues: 856-1330, 'G', 1332-1390, 'KK', 1393-1607, 'RA', 1610-1940 <BOB>
A:Cross-references: EMBL:X51593; NID:g29463; PIDN:CAA35942.1; PID:g29464
A:Experimental source: clone gTMC-E
R:Bober, E.; Buchberger-Seidl, A.; Braun, T.; Singh, S.; Goedde, H.W.; Arnold, H.H.
Eur. J. Biochem. 189, 55-65, 1990
A:Title: Identification of three developmentally controlled isoforms of human myosin
A:Reference number: S09331; MUID:90235862; PMID:1691980
A:Accession: S09333
A:Molecule type: mRNA
A:Residues: 856-901, 'X', 903-971, 'X', 973-1041, 'X', 1043-1111, 'X', 1113-1181, 'X', 1183-125
1, 'X', 1673-1741, 'X', 1743-1811, 'X', 1813-1881, 'X', 1883-1940 <BOB>
A:Cross-references: EMBL:X51593
C:Genetics: GDB:MYH3
A:Gene: GDB:MYH3
A:Cross-references: GDB:119443; OMIM:160720
A:Map position: 17p13.1-17p13.1
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; muscle
F:89-767/Domain: myosin motor domain homology <MMOT>
F:179-186/Region: nucleotide-binding motif A (P-loop)
F:549-586/Region: actin binding #status predicted
F:656-678/Region: actin binding #status predicted
F:840-1940/Domain: coiled coil #status predicted
F:840-1280/Region: S2
F:1281-1940/Region: light meromyosin
F:130/Modified site: N6,N6, trimethyllysine (Lys) #status predicted
F:185/Binding site: ATP (Lys) #status predicted
F:766/Active site: Cys #status predicted

Query Match 12.8%; Score 397; DB 1; Length 1940;
Best Local Similarity 21.5%; Pred. No. 9.2e-07;
Matches 165; Conservative 143; Mismatches 258; Indels 200; Gaps 22;

QY 4 LSLELMKLRNKRTRKMRSMVMYKQEGMELKQATQKDLTSKGIQVLEGKLVSIKREID 63
Db 833 LFFKIKPLKLSAETKEKATMKEE----FQTKDELAKSEAKRKELEKLVTLVQEKND 887
QY 64 ----EKCTEKLELYIQEISASDOVEKCV-----DIAQLEEDLKEKREILSLKQS 112
Db 888 LQLQVQAESENLDD--AEERC--DOLIKAFQLEAKIKETVRADEBEINAEALTAKRK 943
QY 113 LEENIT-FSKOIEDLVTKCOLLETERDNLVSKDRERAETLSAEMOILTERLALEREYK 171
Db 944 LEDECSLEKLDIDDLLETLAKVEKHEKHAENKVNKLTELSELG----LDETAKLTREKA 999
QY 172 LQKELQSQSLQOKE-----LSARLOQOLCSFQEMTSEKNV-----FKEE 214
Db 1000 LQEAHQALDLOAEEDKVNLSLNTKSKLEQOVEDLESSLEQEKRLVLDLERNKRKLEGD 1059
QY 215 LKLALAE-LDAVQKQEESESLRVK-----QLE---EERKSTAEQTLRLDNLRLKEVEL 264
Db 1060 LKLAQESTLDELNDKQDLERLKKDFYQCLQSKVEDEQTLGLQFQKKIKELQARIEL 1119
QY 265 EKHTAAHAQAAILIAQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESQEK-K 323
Db 1170 EEEAEARATRAKTEKQKRSYARELEELSERLEE-----AGVTSTQIELNKKRAE 1171

Qy 324 YNDTAQSLRDVTAQLESE-----QEKYNDTAQSLRDVTAQLESVOEKYNDTAQSLRDVSAQ 379
Db 1172 FLKLRRDLLEATLQHEAWATLRKKHADSVAEALGEQIDNLRVKOK-----1217
Qy 380 LESYKSTLKEIEDLKLENLTQEKVMAEAKSVEDVQOQIITAESTNOEYARMYQDQNR 439
Db 1218 LEKESEFKLEIDDLSSMSVSKANLEKICHTLEDQLSEARGKNBEIORSLSLTQTQ 1277
Qy 440 STLKEEIKETT-----SSFLEKITDLKNLROODEDFRKQLEKGR-----482
Db 1278 KSLQTEAGLSRLQBEKESIVSSLSKQAFQTQTEELKRLQLEENKAKNALAHALQSS 1337
Qy 483 -----TAEKENTMTLMEINKWRLY-----EELYEK 510
Db 1338 RHCDDLREYEEBQGAELQRLSKANSEVAQWRKYETDAIQTRELEEAQEKLAQR 1397
Qy 511 TKPQQQLDAF-----521
Db 1398 LODSEEQEAVNAKASLEKTKQRLQGEVEDLMVDVERANSALAAADKKQRNFKVLAEW 1457
Qy 522 -----EAKOALLNEHGATQOELNKRDSYAOLIG-----HONLKOKIKHVYK 566
Db 1458 KTKCEQAELASLESRSLSLSTELFKLNAYEALDQLETYVRENKNLEGEIADLTQEI 1517
Qy 567 DENSQKSEVSKLSQSLVKRK-----QNELRLQGEIDKALGTR 604
Db 1518 AENGKTIHELEKSRKQIELEKADIQALAEAEAALEHEEAKILRTQ 1563

RESULT 12

148175
A:Title: myosin heavy chain alpha, cardiac muscle [similarity] - golden hamster
C:Species: Mesocricetus auratus (golden hamster)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 02-Feb-2001
C:Accession: I48175; A23938
R:Wang, R.; Sole, M.J.; Cukerman, E.; Liew, C.C.
J. Mol. Cell. Cardiol. 26, 1155-1165, 1994
A:Title: Characterization and nucleotide sequence of the cardiac alpha-myosin heavy chain
A:Reference number: I48153; MUID:95115033; PMID:7815459
A:Accession: I48175
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1939 <RES>
A:Cross-references: GB:I15351; NID:g402373; PIDN:AAB59701.1; PID:g402374
R:Liew, C.C.; Jandreski, M.A.
Proc. Natl. Acad. Sci. U.S.A. 83, 3175-3179, 1986
A:Title: Construction and characterization of the alpha-form of a cardiac myosin heavy chain
A:Reference number: A23938; MUID:86205859; PMID:3458174
A:Accession: A23938
A:Molecule type: mRNA
A:Residues: 1630-1843, 'R', 1845-1878, 'T', 1880-1927, 'N', 1929-1932, 1934-1939 <LIE>
C:Genetics:
A:Introns: 67/3; 115/3; 168/1; 177/2; 214/3; 245/3; 267/1; 300/1; 334/3; 381/1; 470/3; 532/2; 1453/3; 1509/3; 1550/3; 1653/3; 1721/3; 1763/3; 1855/3; 1887/3; 1932/3
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: Atp; cardiac muscle; heart; nucleotide binding; P-loop
F:88-768/Domain: myosin motor domain homology <MMOT>
F:178-185/Region: nucleotide-binding motif A (P-loop)

Query Match 12.7%; Score 393; DB 2; Length 1939;
Best Local Similarity 23.8%; Pred. No. 1.3e-06;
Matches 181; Conservative 132; Mismatches 237; Indels 210; Gaps 30;

Qy 9 MKLRNKRTRKMSVMVQKE--GMELKIQATQKDLTSTESGKIVQLEGKLVSTSEKIDKDC 66
Db 832 MKLYFKIKPLLKSTETEKEMANRDEFGVRVSEKSEARRKELEEKVSLIQEKNDLQF 891
Qy 67 ETEKLEVIQETSCASDOVEKCKVDIAOLE-----EDLKEKREILSLKQSLLEN 116
Db 892 QVQAQMDNDAERCDQLINKKI---QLEAKVEMTERLEDEEEMNAELTSKRRKLSDE 948
Qy 117 IT-FSKQIEDTVKCOLLETEDRNILVSKDRERAETLSAEMQILTERLALEQVEKLOQK 175

Db 949 CSELKOIDDLLETLAKVEKEKHATENK-----VKNLTBEAMAGLDEITIAKTREKALQEA 1004
Qy 176 ELQSOSLLOQKE-----LSARLOOOLCSFOEBMTSEK-----NVF 211
Db 1005 HQQALDDLQABEDKVTITKSVKLEQVDDLEGSLEQEKVKRMDLERAKRKLGGDLNVT 1064
Qy 212 KE-----ELKLALAEALDAVQO--KEEQSERIVKQLEEBRKSTABOLTRLDNL 257
Db 1065 QESIMDLNDKLOLEELKKKEFFDISQNSKIEDQALALQOKKLK--ENQARIEEL- 1120
Qy 258 REKEVELEKHHTAAHAQAILIAQEKYNDTAQSLRDVTAQLESVOEKYNDTAQSLRDVTAQL 317
Db 1121 -EELEAERTARAKVEKL-----RSDLTRELEETISERLEEAGGA--TSVQI 1163
Qy 318 ESEQKYNDTAQSLRDVTAQLESBOEKYNDTAQSLRDVTAQLESVOEKYNDTAQSLRDVS 377
Db 1164 EMNKKREAEFGKMRD-----LEEATLQHEATAAALRKKHA--DSVAE-LGEQIDNLRVK 1216
Qy 378 AQLESYKSSSTLKEIEDL-----KLENL--TLQE-----KVMAEAKSVEDVQ 416
Db 1217 QKLEKESEFKLEDDVTSNMEQIIKAKANLEKVSRTLEDQANEYRVKLEESORSINDFT 1276
Qy 417 QOITAESTNOEYARMYQDQNRSTLKEEIKETIS---SFLEKITDLKNLQRLQO-----468
Db 1277 TORAKLOTENGELARQLEE-----KEALISQITRGLKSYTQOQMEDLKQLEEBGKAKN 1329
Qy 469 -----DED-FRKQLEKGRKTAENKENVMTLMEINKWRLY-----EELY 508
Db 1330 ALAHALQSGARHDCDLLREQYEEEMAKAELQRLVLSKANSEVAQWRKYETDAIQTTELE 1389
Qy 509 EKTTPFOQL-DAFEA-----EKOALINE-----HGATQEQELKIRD 544
Db 1390 EAKKKLAQLQDAEEAVNAKCSSLEKTRHRLQNEIEDLMVDVERSNAAAALDKKOR 1449
Qy 545 SYAQLLGHQNLK-----QKTKHVVKUKDENSOLKS 574
Db 1450 NFDKILAEWKQYEBSESSQSEARSLSLSTELFKLNAYEESLEHLETKRENKNLQE 1509
Qy 575 EVSKLRSOLVK-----RKQNE---LRLOGLDKA 600
Db 1510 EISDLTEQLGEGGKNVHELEKVRKOLEVEKMELOQALBEA 1549

RESULT 13

T14156
A:Title: kinesin-related protein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 11-May-2000
C:Accession: T14156
R:Wood, K.W.; Sakowicz, R.; Goldstein, L.S.; Cleveland, D.W.
Cell 91, 357-366, 1997
A:Title: CNRP-E is a plus end-directed kinetochore motor required for metaphase chromosome segregation
A:Reference number: I17893; MUID:98028574; PMID:9363944
A:Accession: T14156
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2954 <WOO>
A:Cross-references: EMBL:AF027728; NID:g2586070; PID:g2586071; PIDN:AAC60300.1
C:Genetics:
A:Gene: XCENP-E
C:Superfamily: centromere protein E; kinesin motor domain homology

Query Match 12.6%; Score 391; DB 2; Length 2954;
Best Local Similarity 22.5%; Pred. No. 2.3e-06;
Matches 164; Conservative 146; Mismatches 236; Indels 182; Gaps 28;

Qy 11 LRNKRTRKMSVMVQKEG-----ELKIQATQKDLTSTESGKIV-----48
Db 1543 LQKLESLSLENILIKENIDITTLKHHSDTQAOQLQTOQLAKNLATAASDNCPIQEK 1602
Qy 49 -----QLEKGVISIEKEIDKCEKTEKLEYIQEISCASQDVE-RCKVD-----91
Db 1603 ETSADCVHPLEKILLTEELHOKTNEQEKLLHENELEQA--QVELACEVEHLMKSMIE 1660

R:Yamauchi-Takahara, K.; Sole, M.J.; Liew, J.; Ing, D.; Liew, C.C.
Proc. Natl. Acad. Sci. U.S.A. 86, 3504-3508, 1989
A:Title: Characterization of human cardiac myosin heavy chain genes.
A:Reference number: A94224; MUID:89264452; PMID:2726733
A:Accession: A94224
A:Molecule type: DNA
A:Residues: 1-87, 'Q', 89-106, 'E', 108-177, 1325-1702, 'DR', 1705-1786, 1788-1803, 'E', 1804-1935
R:Yamauchi-Takahara, K.; Sole, M.J.; Liew, J.; Ing, D.; Liew, C.C.
Proc. Natl. Acad. Sci. U.S.A. 86, 7416-7417, 1989
A:Reference number: A94226
A:Contents: annotation; erratum
R:Kurabayashi, M.; Tsuchimochi, H.; Komuro, I.; Takaku, F.; Yazaki, Y.
J. Clin. Invest. 82, 524-531, 1988
A:Title: Molecular cloning and characterization of human cardiac alpha- and beta-form myosin atrium.
A:Reference number: A92770; MUID:88299163; PMID:2969919
A:Accession: B28908
A:Molecule type: mRNA
A:Residues: 1412-1518, 'R', 1520-1574, 'NV', 1577-1935 <KOR>
A:Cross-references: GB:M21665
R:Notice: the authors translated the codon AGC for residue 108 as Arg
R:Lichter, P.; Umada, P.K.; Levin, J.E.; Vosberg, H.P.
Eur. J. Biochem. 160, 419-426, 1986
A:Title: Partial characterization of the human beta-myosin heavy-chain gene which is expressed in adult human atrium.
A:Reference number: A24997; MUID:87030293; PMID:3021460
A:Accession: A24997
A:Molecule type: DNA
A:Residues: 682-721, 975-1112, 1854-1935 <LIC>
A:Cross-references: GB:X04627
R:Saez, L.J.; Gnanola, K.M.; McNally, E.M.; Peghali, R.; Eddy, R.; Shows, T.B.; Leinwand
Nucleic Acids Res. 15, 5443-5459, 1987
A:Title: Human cardiac myosin heavy chain genes and their linkage in the genome.
A:Reference number: A93669; MUID:87260010; PMID:3037493
A:Accession: A27858
A:Molecule type: DNA
A:Residues: 1854-1865, 'A', 1867-1935 <SAE>
A:Cross-references: GB:X05631; GB:Y00362; NID:g34643; PIDN:CAA29119.1; PID:g34644
R:Diederich, K.W.; Eisele, I.; Ried, T.; Jaenicke, T.; Lichter, P.; Vosberg, H.P.
Hum. Genet. 81, 214-220, 1989
A:Title: Isolation and characterization of the complete human beta-myosin heavy chain gene.
A:Reference number: 154254; MUID:89154425; PMID:2522082
A:Accession: 154254
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 653-720 <RES>
A:Cross-references: GB:M27636; NID:g179511; PIDN:AAA79019.1; PID:g601916
R:Bober, E.
submitted to the EMBL Data Library, January 1989
A:Reference number: S12458
A:Accession: S12458
A:Molecule type: mRNA
A:Residues: 785-1076, 'E', 1078-1123, 'A', 1125-1702, 'DE', 1705-1935 <BOB>
A:Cross-references: EMBL:X51591; NID:g29467; PIDN:CAA35940.1; PID:g29468
R:Bober, E.; Buchberger-Seidl, A.; Braun, T.; Singh, S.; Goedde, H.W.; Arnold, H.H.
Eur. J. Biochem. 189, 55-65, 1990
A:Title: Identification of three developmentally controlled isoforms of human myosin heavy chain.
A:Reference number: S09331; MUID:90235862; PMID:1691980
A:Accession: S09331
A:Molecule type: mRNA
A:Residues: 785-830, 'X', 832-900, 'X', 902-970, 'X', 972-1040, 'X', 1042-1076, 'E', 1078-1110, 'X', 1602-1670, 'X', 1672-1702, 'DE', 1705-1740, 'X', 1742-1810, 'X', 1812-1935 <BOB>
A:Cross-references: EMBL:X51591
R:Jandreski, M.A.; Liew, C.C.
Hum. Genet. 76, 47-53, 1987
A:Title: Construction of a human ventricular cDNA library and characterization of a beta myosin heavy chain.
A:Reference number: S02229; MUID:8719738; PMID:3032769
A:Accession: S02229
A:Molecule type: mRNA
A:Residues: 1393-1702, 'DR', 1705-1935 <JAN>
A:Cross-references: EMBL:X06976; NID:g34860; PIDN:CAA30039.1; PID:g825694
R:Saez, L.; Leinwand, L.A.
Nucleic Acids Res. 14, 2951-2969, 1986
A:Title: Characterization of diverse forms of myosin heavy chain expressed in adult human

A:Reference number: A93616; MUID:86176778; PMID:2421254
A:Accession: B23767
A:Molecule type: mRNA
A:Residues: 'LLGVELASG', 1311-1312, 'G', 1314-1355, 'R', 1357-1358, 'CD', 1361-1438, 'LQ', 14
A:Note: the first ten codons of the sequence figure show the reverse complementary strand.
C:Genetics:
A:Gene: GDB:MYH7
A:Cross-references: GDB:120215; OMIM:160760
A:Map position: 14q12-14q12
A:Introns: 67/3; 115/3; 168/1; 177/2; 213/3; 244/3; 266/1; 299/1; 333/3; 380/1; 419/3;
24/3; 1390/2; 1451/3; 1507/1; 1548/3; 1651/3; 1719/3; 1761/3; 1853/3; 1885/3; 1930/3
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: actin binding; ATP; cardiac muscle; coiled coil; heart; hydrolase; methyl
F:88-765/Domain: myosin motor domain homology <MMOT>
F:178-185/Region: nucleotide-binding motif A (P-loop)
F:548-585/Region: actin binding #status predicted
F:655-677/Region: actin binding #status predicted
F:839-1935/Domain: coiled coil #status predicted <COI>
F:839-1279/Region: S2
F:1280-1935/Region: light meromyosin
F:129/Modified site: N6, N6, trimethyllysine (Lys) #status predicted
F:184/Binding site: ATP (Lys) #status predicted
F:695,705/Active site: Cys #status predicted

Query Match 12.5%; Score 389.5; DB 1; Length 1935;
Best Local Similarity 23.4%; Pred. No. 1.7e-06;
Matches 175; Conservative 131; Mismatches 238; Indels 203; Gaps 30;

Qy 14 KREYKMRSMVQEGMEKLQATQKDLTSGKIVQLSGKLVSIEKEIDEKCETEKLE 73
Db 844 BREKEMASW--KEEFTRLK-----BALEKSEARRKELEKVMYSLQEQKNDLQVQVQAD 896
Qy 74 YIQETSCASDQVEKCKVDIAOLE-----EDLKEKDRILSLKSLLENIT--FSKQ 122
Db 897 NLADAEKCDQLINKKI---OLEAKVKEMNERLEDEEMNALTAKKRLKEDECSLAK 953
Qy 123 IEDTVKCOLLETEDRNLVSKDRRAETLSAEMQTLTSLALERQYKQLQKELQSQSL 182
Db 954 IDDLTLAKVEKEKHAENK---VKNLTDEMAGLDEIAKLTREKKAQEAHQALDD 1009
Qy 183 LQOEKE-----LSARLQQLCSQEQEMTSEKNV-----FKELKALAE-LDA 224
Db 1010 LQAEEDKYNTLTAKVKLEQQVDDLEGSLEQEKVMDLRAKRLKLEGLDKLTQESIMDL 1069
Qy 225 VQKEEQSERLVK-----OLEEER---KSTAEQLTRLDNLLRKEVELEKHAH 271
Db 1070 ENDKQDLERLKKDFELNALNARDEQALGSQKQKLEQARIEELEEESEPERTA 1129
Qy 272 AQAILIAQEKYNDTAQSLR-DVTAQLESVQEKYNDTAQSLRDVTAQLESQEKYNDTAQS 330
Db 1130 AK-----VEKLRSDLSRELEISERLEAGGA---TSVQIEMNKKREAEFFQKM 1174
Qy 331 LRDVTAQLESQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESQEKYNDTAQS 390
Db 1175 RRD-----LEEATLQEAATAALRKHA--DSVAE--LGEQIDNLQVRVKQLEKESEFKLE 1227
Qy 391 IEDL-----KLENL--TLQE-----KVANAESVEDVQOQILTAESTNOEY 429
Db 1228 LDDVTNNMEQIIKANKLEKMTLEDQNMHRSKAEETQSVNDLTQAKLQTEGEL 1287
Qy 430 ARMVQDLQNRSTLKPEEIKETIS--SFLEKTTDLKNQLRQO-----DE 470
Db 1288 SRQLD-----KEALISQLTRGLTYTQOEDLKRQLEEEVKAKNALAHQASRHC 1340
Qy 471 D-FRKQLEEKGRTRAEKNVMTLMEINWKRLY-----EELYEKTRPFQOOL--- 518
Db 1341 DLLREQYEEETEAKAELQRLVLSKANSEVAQWRKTYKTDALQIRTEELEBAKKKLAQLOEA 1400
Qy 519 -----DAFAEAKQALINE-----HGATQEQNLKTRDSYAQLLGHQNLK- 556
Db 1401 EEAEEAVNAKSSLEKTKHRLQNEIEDLMVDVRSNAAAALDKKQRFKILAEWKQY 1460
Qy 557 -----QKIKHVVKLDKNSQLKSEVSKLSQLVK-- 585

Db 1461 EESQSESSQKPARSLSTELFKLNAYEESLEHLETFKRENKNLQOEISDLTEQLGSSG 1520
QY 586 -----RKONE---LRLQGLDKA 600
Db 1521 KTIHELEKVRKQLEAEKMELOSALEA 1547

Search completed: December 20, 2002, 15:12:55
Job time : 34.2176 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:06:33 ; Search time 32.5737 Seconds
(without alignments)
2581.255 Million cell updates/sec

Title: US-09-685-010-48
Perfect score: 3104
Sequence: 1 MRALSFLMKLRNKRKTKMR.....FCHASKENFTPLKEGPNCC 631

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3104	100.0	631	ABG60843	Mouse receptor for
2	3104	100.0	794	AAU11437	Mouse hyaluronic a
3	3098	99.8	631	AAU11437	RHAMM 1-2a isoform
4	3086.5	99.4	630	AAW39166	Mouse RHAMM protei
5	2952.5	95.1	606	AAW39167	Receptor for hyalu
6	2374.5	76.5	713	AAU11438	Rat hyaluronic aci
7	2295.5	74.0	476	AAU11438	Hyaluronan recepto
8	1953	62.9	725	AAW39165	Human RHAMM protei
9	1953	62.9	725	ABG60842	Human receptor for
10	1952	62.9	725	AAU11436	Human hyaluronic a

11	1205	38.8	351	17	AAW01052	Human umbilical ve
12	421.5	13.6	1948	22	ABG21233	Nucle human diagno
13	414	13.3	1411	17	ABW02258	Nucleolar/endosoma
14	396.5	12.8	1690	22	ABB61144	Drosophila melanog
15	396.5	12.8	1690	22	ABB61173	Drosophila melanog
16	391	12.6	2954	20	AAW01632	Amino acid sequenc
17	388.5	12.5	1939	23	ABB77096	Human alpha-myosin
18	387	12.5	1886	19	AAW54241	Rattus norvegicus
19	386	12.4	1392	20	AAW06999	Retin protein seq
20	384	12.4	1960	22	AAW78854	Human protein seq
21	384	12.4	1979	21	AAW18171	Plasmodium falcipa
22	384	12.4	2143	22	ABG01716	Novel human diagno
23	383	12.3	1427	12	AAW10534	Human 160kD mediat
24	380.5	12.3	1963	22	AAW79838	Human protein seq
25	375.5	12.1	1988	22	AAW40999	Human polypeptide
26	375.5	12.1	1988	22	AAW41000	Human polypeptide
27	375	12.1	1388	23	AAW79590	Human kinesin moto
28	375	12.1	1388	23	AAE14400	Human kinesin supe
29	374	12.0	1489	22	ABB59948	Drosophila melanog
30	370.5	11.9	931	22	AAW79504	Human protein seq
31	370.5	11.9	990	22	AAW78520	Human protein seq
32	363	11.7	1857	23	AAU84350	Protein MYH11 diff
33	362	11.7	795	23	ABB77430	Human tumour marke
34	362	11.7	885	16	AAW66930	AMML chromosome in
35	362	11.7	976	22	AAG66581	Human SCP-1 mutein
36	362	11.7	2056	22	ABB59344	Human SCP-1
37	360.5	11.6	1354	18	AAW23654	Drosophila melanog
38	360.5	11.6	1354	19	AAW71020	Physiologically ac
39	360.5	11.6	1354	20	AAW07082	A modified rho tar
40	360	11.6	2383	23	ABG05631	Renal cancer assoc
41	358	11.5	2117	22	AAU32040	Human breast speci
42	358	11.5	2192	18	AAW21732	Novel human secret
43	358	11.5	2272	18	AAW21731	LexA/NuMA fusion p
44	358	11.5	3248	17	AAW99795	GAL4/HA/NuMA fusio
45	357.5	11.5	2101	15	AAW47173	Kinetochoe protei
						Sequence of the in

ALIGNMENTS

RESULT 1
ABG60843
ID ABG60843 standard; Protein; 631 AA.
XX
AC ABG60843;
XX
DT 13-AUG-2002 (first entry)
XX
DE Mouse receptor for hyalauroan-mediated motility (RHAMM).
XX

Tissue disorder; response-to-injury process; cell proliferating;
hyaluronic acid; HA; receptor for hyalauroan-mediated motility;
RHAMM; inflammatory neurological disorder; Parkinson's disease;
Kw Alzheimer's disease; arthritis; multiple sclerosis; gastritis; nephritis;
Kw inflammatory dermatosis; psoriasis; inflammatory bowel disease;
Kw stenosis; restenosis; cancer; kidney fibrosis; inflammatory lung disease;
Kw emphysema; asthma; cystic fibrosis; obesity; obesity related disease;
Kw lupus; cardiovascular disease; atherosclerosis; wound; scar; diabetes;
Kw tissue transplantation; stroke; inflammatory response; fibrotic response;
Kw medical implant; Acquired immunodeficiency syndrome; AIDS; hepatitis;
Kw myocardial fibrosis; hepatic fibrosis; chronic cystitis; acute mastitis;
Kw septic shock; thyroiditis; retinopathy.

XX Mus musculus.
OS
XX WO200228415-A1.
PN
XX 11-APR-2002.
PD
XX 05-OCT-2000; 2000WO-IB01534.
PF
XX 05-OCT-2000; 2000WO-IB01534.
PR
XX

PA (TRAN-) TRANSITION THERAPEUTICS & DIAGNOSTICS IN.

PI Turley EA, Cruz TF;

XX WPI; 2002-435298/46.

XX Treating tissue disorder associated with response-to-injury process or
PT proliferating cells in mammals, e.g. fibrosis, inflammation, by
PT administering a compound that alters activity of transition molecules
PT within a cell

XX Disclosure; Fig 50; 215pp; English.

XX The invention describes a method of treating a tissue disorder associated
CC with response-to-injury process or proliferating cells in a patient,
CC comprising administering a polypeptide (I) which binds hyaluronic acid
CC (HA), an antibody which binds one of domains DI-D5 of Receptor for
CC hyaluronan-mediated motility (RHAMM), a polypeptide fragment encoding
CC any of DI-D5 of RHAMM, or a vector which expresses antisense RHAMM,
CC antibodies or a polypeptide fragment. The method is useful for treating a
CC patient with an inflammatory neurological disorder such as Parkinson's
CC disease, Alzheimer's disease, arthritis including rheumatoid arthritis,
CC osteoarthritis, multiple sclerosis, inflammatory dermatosis (psoriasis),
CC inflammatory bowel disease, stenosis or restenosis, cancer, kidney
CC fibrosis, inflammatory lung disease (e.g. emphysema, asthma, cystic
CC disease), obesity or obesity related diseases, lupus, cardiovascular
CC disease (e.g. atherosclerosis), and wound especially surgical excision
CC adhesions, to prevent scar and also for treating or preventing diabetes
CC mellitus. The method is also useful for treating tissue transplantation
CC (e.g. skin grafts), stroke, inflammatory responses or fibrotic response
CC associated with medical implants such as hip implants, vascular wraps and
CC catheters), inflammatory diseases such as AIDS, myocardial and hepatic
CC fibrosis, chronic cystitis, acute mastitis, gastritis, nephritis,
CC hepatitis, septic shock, thyroiditis, and retinopathy. This sequence
CC represents a receptor for hyaluronan-mediated motility protein used in
CC the method of treating a tissue disorder described in the invention.

XX SQ Sequence 631 AA;

Query Match 100.0%; Score 3104; DB 23; Length 631;
Best Local Similarity 100.0%; Pred. No. 6.1e-153;
Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALSLEMLKRNKRETKMRSMVMQEGMELKQATQKDLTSKGIQVLEGKLSIEKE 60

DB 1 MRALSLEMLKRNKRETKMRSMVMQEGMELKQATQKDLTSKGIQVLEGKLSIEKE 60

QY 61 KIDKCEKTEKILEYQIEISASDQVEKCKVDIAQLEEDLKEKDRILSLKQSLLENITFS 120

DB 61 KIDKCEKTEKILEYQIEISASDQVEKCKVDIAQLEEDLKEKDRILSLKQSLLENITFS 120

QY 121 KQIEDLTVMKQQLLETERDNLVSKDRERATLSAEMQILTERLALEROEYKLOQKELQSO 180

DB 121 KQIEDLTVMKQQLLETERDNLVSKDRERATLSAEMQILTERLALEROEYKLOQKELQSO 180

QY 181 SLLOQEKELSLARLQOLCSFQEBMTSEKNVFKELKLALAEALDAVQKQEEQSERLVKQLE 240

DB 181 SLLOQEKELSLARLQOLCSFQEBMTSEKNVFKELKLALAEALDAVQKQEEQSERLVKQLE 240

QY 241 EERKSTAEQTLRLDNLRLKEVELEKHAHAQAAILIAQEKYNDTAQSLRDVTAQLESVQ 300

DB 241 EERKSTAEQTLRLDNLRLKEVELEKHAHAQAAILIAQEKYNDTAQSLRDVTAQLESVQ 300

QY 301 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 360

DB 301 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 360

QY 361 SVQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 420

DB 361 SVQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 420

QY 421 TAESTNQEVARMVQDLQNRSTLKEEIKETISFLEKITDNLQNRQOQDEDFRKQLEEK 480

DB 421 TAESTNQEVARMVQDLQNRSTLKEEIKETISFLEKITDNLQNRQOQDEDFRKQLEEK 480

DB 421 TAESTNQEVARMVQDLQNRSTLKEEIKETISFLEKITDNLQNRQOQDEDFRKQLEEK 480

QY 481 KRTAEKENVMTELTMEINKWRLLYEELYEELKTPFQOQLDAFAEAKQALLNEHGATQEQLN 540

DB 481 KRTAEKENVMTELTMEINKWRLLYEELYEELKTPFQOQLDAFAEAKQALLNEHGATQEQLN 540

QY 541 KTRDSYAQLLGHONLKQKTKHVVKLKDENSQKSEVSKLRSQLVKRRQKQNELRLQGEELDKA 600

DB 541 KTRDSYAQLLGHONLKQKTKHVVKLKDENSQKSEVSKLRSQLVKRRQKQNELRLQGEELDKA 600

QY 601 LGIRHFDPSKAFCHASKENFTPLKSGNPNC 631

DB 601 LGIRHFDPSKAFCHASKENFTPLKSGNPNC 631

RESULT 2

AAU11437

ID AAU11437 standard; Protein; 794 AA.

AC AAU11437;

DT 12-MAR-2002 (first entry)

XX Mouse hyaluronic acid binding protein RHAMM.

DE Mouse; hyaluronic acid binding protein; RHAMM; gene therapy;

KW receptor for HA mediated motility; immunosuppressive; cytostatic.

KW conjugate; rheumatoid arthritis; scleroderma; liver fibrosis; cancer;

OS Mus musculus.

FX Key Location/Qualifiers

FT Misc-difference 55 /note= "Encoded by AAC"

FT Misc-difference 71 /note= "Encoded by ACG"

FT Misc-difference 89 /note= "Encoded by CAA"

FT Misc-difference 91 /note= "Encoded by CAC"

FT Misc-difference 540 /note= "Encoded by ACT"

FT Misc-difference 668 /note= "Encoded by GAT"

PN WO200180899-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-CA00533.

XX 20-APR-2000; 2000US-198613P.

XX (CANG-) CANGENE CORP.

PI Woloski BMR, Williams AM, Sereda TJ, Wiebe DJ;

XX WPI; 2002-075094/10.

DR N-PSDB; AAS17497.

XX Protein conjugates that selectively target certain tissues and organs
PT useful for treating and preventing various diseases, comprises
PT glucose-aminoglycan-targeting domain conjugated to a therapeutic
PT protein

PS Claim 6; Page 116; 121pp; English.

XX The invention relates to a conjugate comprising an hyaluronic acid (HA)
CC -binding protein e.g. RHAMM (receptor of HA mediated motility) or peptide
CC contiguous with, or coupled to a polypeptide conjugated to a therapeutic
CC agent, and the polynucleotides encoding them. Also included is a method
CC for preparation of the HA-binding protein by inserting a first nucleotide
CC sequence encoding a HA-binding protein directly linked to a second

CC nucleotide sequence encoding a therapeutic protein into a suitable
CC vector, expressing the vector in an acceptable host, purifying conjugate
CC molecule from host or expression medium. The composition is useful for
CC altering in vivo the distribution of a therapeutic agent comprising
CC administering the composition to the animal where conjugate molecule will
CC distribute primarily in tissues and organs containing high levels of
CC endogenous HA and for treating mammal with a disorder where a diseased
CC tissue of the mammal contains high level of HA e.g. rheumatoid
CC arthritis, scleroderma, liver fibrosis and cancer. Lower therapeutic
CC dosages required also translates into lower immunogenicity of the
CC conjugated protein as compared to the native protein. As a result,
CC conjugates improve patient compliance and reduce direct and indirect
CC costs associated with the drug substance and its administration.
CC Conjugates allows for the use, where appropriate, of lower, safer,
CC dosages as compared to the conventional dosage requirements for the
CC unconjugated corresponding therapeutic agent. Conjugate molecules have an
CC increased half-life and potency, resulting in prolonged circulation of
CC the molecule, efficient distribution into the target tissues, and
CC increased bioavailability. The present sequence represents a RHAMM
CC protein.
XX
XX

SQ Sequence 794 AA;
Query Match 100.0%; Score 3104; DB 23; Length 794;
Best Local Similarity 100.0%; Pred. No. 7.8e-153;
Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRALSLEMLKLNKRETKMRSMVMVQEGMELKLTQKDLTSGKIVQLGKLVSIKE 60
Db 164 MRALSLEMLKLNKRETKMRSMVMVQEGMELKLTQKDLTSGKIVQLGKLVSIKE 223
Qy 61 KIDECETEKLEYIQEISCDASQDVCKVKVDALEEDLKEKDRILSLKQSLSENIITFS 120
Db 224 KIDECETEKLEYIQEISCDASQDVCKVKVDALEEDLKEKDRILSLKQSLSENIITFS 283
Qy 121 KOIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALERQYKELQKELQSQ 180
Db 284 KOIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALERQYKELQKELQSQ 343
Qy 181 SLQQEKELSLARLQQLCSFQEMTSKRVKPEELKLAELDAVQOKEEQSERLVKQLE 240
Db 344 SLQQEKELSLARLQQLCSFQEMTSKRVKPEELKLAELDAVQOKEEQSERLVKQLE 403
Qy 241 EERKSTAQRLTDLNLLREKEVELEKHAHAQAQILIAQEKYNDTAQSLRDVTAQLESVQ 300
Db 404 EERKSTAQRLTDLNLLREKEVELEKHAHAQAQILIAQEKYNDTAQSLRDVTAQLESVQ 463
Qy 301 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 360
Db 464 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 523
Qy 361 SVQEKYNDTAQSLRDVTAQLESYKSTLKEIEDLKLNLTLQEKVAMAEKSVEDVQQQIL 420
Db 524 SVQEKYNDTAQSLRDVTAQLESYKSTLKEIEDLKLNLTLQEKVAMAEKSVEDVQQQIL 583
Qy 421 TAESTNQEYARMVDQLNKSTLKEIEKITSFLEKITDLKNLQROQDEDFRQLEBK 480
Db 584 TAESTNQEYARMVDQLNKSTLKEIEKITSFLEKITDLKNLQROQDEDFRQLEBK 643
Qy 481 KRTAEKENVMTLMEINKRLLYBELYEKTKPQQQDLDAFAEAKQALLNEHGATQEQLN 540
Db 644 KRTAEKENVMTLMEINKRLLYBELYEKTKPQQQDLDAFAEAKQALLNEHGATQEQLN 703
Qy 541 KIRDSYAQLLGHQNLKQIKHVKVLDKNSQLKSVSKLRSQLVKRNQNEURLQGLDKA 600
Db 704 KIRDSYAQLLGHQNLKQIKHVKVLDKNSQLKSVSKLRSQLVKRNQNEURLQGLDKA 763
Qy 601 LGIRHFDPSKAFCHASKENFTPLKEGNPNC 631
Db 764 LGIRHFDPSKAFCHASKENFTPLKEGNPNC 794

RESULT 3

AAR99675
ID AAR99675 standard; Protein; 631 AA.
XX
AC AAR99675;
XX
DT 10-OCT-1996 (first entry)
XX
DE RHAMM 1-2a isoform.
XX
KW RHAMM 1-2a; receptor for hyaluronic acid mediated motility;
KW hyaluronan receptor; cell locomotion; cell proliferation;
KW breast cancer; therapy.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT Region 55..79
FT /note= "exon 2A-encoded region"
XX
XX EP721012-A2.
XX
XX 10-JUL-1996.
XX
XX 16-OCT-1995; 95EP-0307310.
XX
PR 14-OCT-1994; 94GB-0020740.
XX
XX (MANI-) MANITOBA CANCER TREATMENT & RES FOUND.
XX (UYMA-) UNIV MANITOBA.
XX
PI Entwistle J, Turley EA, Zhang S;
XX
XX WPI; 1996-310997/32.
XX N-PSDB; AAT34525.
XX
PT Receptor for hyaluronic acid-mediated motility protein, and DNA
PT encoding it - useful to treat or prevent diseases associated with
PT the receptor, e.g. breast cancer
XX
PS Claim 8; Page 50-52; 117pp; English.
XX
CC RHAMM 1-2a (AAR99675) is an alternatively spliced variant of
CC RHAMM 1 (AAR99673) (receptor for hyaluronic acid mediated
CC motility), a protein involved in cell locomotion or motility and
CC cell proliferation and transformation. It differs from RHAMM 1
CC by an insertion of 25 amino acids (see also AAR99674) between
CC amino acids 54 and 55 of RHAMM 1, resulting from an alternatively
CC spliced exon 2A (AAT34502). RHAMM 1-2a is the isoform that is
CC overexpressed in tumours. Determination of the level of RHAMM
CC 1-2a in a sample can be used to assess the prognosis of a tumour
CC (esp. breast cancer) patient. The RHAMM 1-2a protein can also
CC be used to suppress or control a tumour by modulating the
CC interaction of cell-associated RHAMM with its ligand.
XX
SQ Sequence 631 AA;

Query Match 99.8%; Score 3098; DB 17; Length 631;
Best Local Similarity 99.7%; Pred. No. 1.3e-152;
Matches 629; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MRALSLEMLKLNKRETKMRSMVMVQEGMELKLTQKDLTSGKIVQLGKLVSIKE 60
Db 1 MRALSLEMLKLNKRETKMRSMVMVQEGMELKLTQKDLTSGKIVQLGKLVSIKE 60
Qy 61 KIDECETEKLEYIQEISCDASQDVCKVKVDALEEDLKEKDRILSLKQSLSENIITFS 120
Db 61 KIDECETEKLEYIQEISCDASQDVCKVKVDALEEDLKEKDRILSLKQSLSENIITFS 120
Qy 121 KOIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALERQYKELQKELQSQ 180
Db 121 KOIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALERQYKELQKELQSQ 180
Qy 181 SLQQEKELSLARLQQLCSFQEMTSKRVKPEELKLAELDAVQOKEEQSERLVKQLE 240

|||||
Db 181 SLLOQEKLSARLQOLCSFQEMTSEKNVFKBELKALAEALDAVQKKEQSERLVKQLE 240
QY 241 EERKSTAEOLTRDLNLLREKEVELEKHIAHAQAIIIAQEKYNDTAQSLRDVTAQLESVQ 300
Db 241 EERKSTAEOLTRDLNLLREKEVELEKHIAHAQAIIIAQEKYNDTAQSLRDVTAQLESVQ 300
QY 301 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 360
Db 301 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 360
QY 361 SVOEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 420
Db 361 SVOEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 420
QY 421 TAESTNOEYARMVODLQNRSTLKEEIKETSSFKETITDLKNLROODEDFRKOLEEK 480
Db 421 TAESTNOEYARMVODLQNRSTLKEEIKETSSFKETITDLKNLROODEDFRKOLEEK 480
QY 481 KRTAEKENVMTELTMEINKWRLLYEELYEKTPFQOQLDAFAEAKQALLNEHGATQEQLN 540
Db 481 KRTAEKENVMTELTMEINKWRLLYEELYEKTPFQOQLDAFAEAKQALLNEHGATQEQLN 540
QY 541 KIRDSYAQLLGHQNLKQKIKHVVKLKDENSQKLSKSEVSKLSRQKQNELRLQGELOKA 600
Db 541 KIRDSYAQLLGHQNLKQKIKHVVKLKDENSQKLSKSEVSKLSRQKQNELRLQGELOKA 600
QY 601 LGIRHFDPSKAFCHASKENFTPLKEGPNCC 631
Db 601 LGIRHFDPSKAFCHASKENFTPLKEGPNCC 631

RESULT 4
AAW39166
ID AAW39166 standard; Protein; 630 AA.
XX
AC AAW39166;
XX
DT 27-APR-1998 (first entry)
XX
DE Mouse RHAMM protein.
XX
KW Hyaluronan receptor; receptor for hyalurononic acid mediated motility;
KW RHAMM; glycosaminoglycan; binding domain; mouse; oncogene; treatment;
KW growth factor; cell locomotion disorder; dementia; detection;
KW inflammatory disorder; autoimmune disease; diagnosis; prognosis.
OS
OS Mus sp.
FH Key Location/Qualifiers
FT Region 279..382
FT Binding-site /note= "repeat region"
FT Binding-site 555..565
FT Binding-site /note= "Hyalurononic acid binding domain"
FT Binding-site 577..586
FT Binding-site /note= "Hyalurononic acid binding domain"
XX
XX WO9738098-A1.
XX
XX 16-OCT-1997.
XX
XX 10-APR-1997; 97WO-CA00240.
XX
XX 10-APR-1996; 96GB-0007441.
XX
XX (MANI-) MANITOBA CANCER TREATMENT & RES FOUND.
XX (UTMA-) UNIV MANITOBA.
XX Entwistle J, Turley EA;
XX WPI; 1997-512715/47.
XX N-PSDB; AAW02801.
XX

Isolated human receptor for hyalurononic acid mediated motility - used
to develop products for treating e.g. tumours, inflammatory
disorders, dementia, AIDS, diabetes and auto-immune diseases
Dislosure; Page 46; 66pp; English.
This sequence represents the mouse hyaluronan receptor which is also
known as the receptor for hyalurononic acid mediated motility (RHAMM).
Hyaluronan is a large glycosaminoglycan that is ubiquitous in the
extracellular matrix and whose synthesis has been linked to cell
migration, growth and transformation. It interacts with cell surfaces via
specific protein receptors, e.g. RHAMM, that mediate many biological
effects. The RHAMM/hyalurononic acid interaction is involved in
oncogene- and growth factor-mediated cell locomotion. The products can be
used in the treatment of disorders involving cell locomotion, e.g. tumour
invasion, birth defects, acute and chronic inflammatory disorders,
Alzheimer's and other forms of dementia, autoimmune diseases, corneal
Huntington's diseases, AIDS, diabetes, including Parkinson's and
dysplasias and hypertrophies, burns, surgical incisions and adhesions,
strokes and multiple sclerosis. They can also be used in e.g. CNS and
spinal cord regeneration, contraception and in vitro fertilisation and
embryo development. The products can also be used in detection, diagnosis
and prognosis.
Query Match 99.4%; Score 3086.5; DB 18; Length 630;
Best Local Similarity 99.7%; Pred No. 4.9e-152;
Matches 629; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
QY 1 MRALSLELMKLRNKRTRMRSMMVQEGMELKQATQKDLTSEKGVQLEKLSIEKE 60
Db 1 MRALSLELMKLRNKRTRMRSMMVQEGMELKQATQKDLTSEKGVQLEKLSIEKE 60
QY 61 KIDECETEKLLLEYIQETSCASDQVEKCKVDIAQLEEDLKEKDRILSKQSEENITPS 120
Db 61 KIDECETEKLLLEYIQETSCASDQVEKCKVDIAQLEEDLKEKDRILSKQSEENITPS 120
QY 121 KOIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALEROEYKQLOQLSQ 180
Db 121 KOIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALEROEYKQLOQLSQ 180
QY 181 SLLOQEKLSARLQOLCSFQEMTSEKNVFKBELKALAEALDAVQKKEQSERLVKQLE 240
Db 181 SLLOQEKLSARLQOLCSFQEMTSEKNVFKBELKALAEALDAVQKKEQSERLVKQLE 240
QY 241 EERKSTAEOLTRDLNLLREKEVELEKHIAHAQAIIIAQEKYNDTAQSLRDVTAQLESVQ 300
Db 241 EERKSTAEOLTRDLNLLREKEVELEKHIAHAQAIIIAQEKYNDTAQSLRDVTAQLESVQ 300
QY 301 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 360
Db 301 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 360
QY 361 SVOEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 420
Db 361 S-QEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 419
QY 421 TAESTNOEYARMVODLQNRSTLKEEIKETSSFKETITDLKNLROODEDFRKOLEEK 480
Db 420 TAESTNOEYARMVODLQNRSTLKEEIKETSSFKETITDLKNLROODEDFRKOLEEK 479
QY 481 KRTAEKENVMTELTMEINKWRLLYEELYEKTPFQOQLDAFAEAKQALLNEHGATQEQLN 540
Db 480 KRTAEKENVMTELTMEINKWRLLYEELYEKTPFQOQLDAFAEAKQALLNEHGATQEQLN 539
QY 541 KIRDSYAQLLGHQNLKQKIKHVVKLKDENSQKLSKSEVSKLSRQKQNELRLQGELOKA 600
Db 540 KIRDSYAQLLGHQNLKQKIKHVVKLKDENSQKLSKSEVSKLSRQKQNELRLQGELOKA 599
QY 601 LGIRHFDPSKAFCHASKENFTPLKEGPNCC 631
Db 600 LGIRHFDPSKAFCHASKENFTPLKEGPNCC 630

RESULT 5
ID AAR99673 standard; Protein; 606 AA.
XX AC AAR99673;
XX 10-OCT-1996 (first entry)
XX DE Receptor for hyaluronic acid mediated motility RHAMM 1.
XX RHAMM 1; receptor for hyaluronic acid mediated motility;
KW hyaluronan receptor; cell locomotion; cell proliferation;
KW breast cancer; therapy.
XX OS Mus sp.
XX FH Key Location/Qualifiers
FT Modified-site 91..93
FT /label= N-glycosylation_site
FT Modified-site 258..260
FT /label= N-glycosylation_site
FT Modified-site 279..281
FT /label= N-glycosylation_site
FT Modified-site 300..302
FT /label= N-glycosylation_site
FT Modified-site 321..323
FT /label= N-glycosylation_site
FT Modified-site 342..344
FT /label= N-glycosylation_site
FT Modified-site 373..375
FT /label= N-glycosylation_site
FT Modified-site 413..415
FT /label= N-glycosylation_site
FT Domain 532..542
FT Domain 553..562
FT /label= Hyaluronan_binding_domain-I
FT Modified-site 594..596
FT /label= N-glycosylation_site
XX EP721012-A2.
XX 10-JUL-1996.
XX PF 16-OCT-1995; 95EP-0307310.
XX 14-OCT-1994; 94GB-0020740.
XX (MANI-) MANITOBA CANCER TREATMENT & RES FOUND.
XX (UYMA-) UNIV MANITOBA.
XX Entwistle J, Turley EA, Zhang S;
XX WPI; 1996-310997/32.
XX N-PSDB; AAT34499;
XX N-PSDB; AAT34500;
XX N-PSDB; AAT23423;
XX N-PSDB; AAT34524.
XX Receptor for hyaluronic acid-mediated motility protein, and DNA
PT encoding it - useful to treat or prevent diseases associated with
PT the receptor, e.g. breast cancer
XX Claim 2; Page 40-42; 117pp; English.
XX RHAMM 1 (AAR99673), or Receptor for Hyaluronic Acid Mediated Motility,
CC is a hyaluronan receptor protein which is involved in cell locomotion
CC or motility and cell proliferation and transformation. Its amino
CC acid sequence was deduced from a cDNA clone (AAT34499) obtd. from
CC murine 3T3 cells and a genomic sequence (AAT34500) from a mouse
CC fibroblast genomic library. 2 Alternative mRNAs for RHAMM 1 were

CC found, RHAMM 1A (see also AAT34523) and RHAMM 1B (AAT34524), that had
CC identical translated portions. Increased expression of RHAMM 1
CC protein is indicative of a poor prognosis for breast cancer. The
CC protein can be used to suppress or control a tumour by modulating
CC the interaction of cell-associated RHAMM with its ligand.
XX SQ Sequence 606 AA;
Query Match 95.1%; Score 2952.5; DB 17; Length 606;
Best Local Similarity 95.7%; Pred. No. 4e-145;
Matches 604; Conservative 2; Mismatches 0; Indels 25; Gaps 1;
Qy 1 MRALSLELMKLRNKRKTRMRSMMVKQEGMELKQATQKDLTSGKGIYQLEGLKLSIEKE 60
Db 1 MRALSLELMKLRNKRKTRMRSMMVKQEGMELKQATQKDLTSGKGIYQLEGLKLSIEKE 54
Qy 61 KIDKCEKTEKLELLEYIOEISCASDQVEKCKVDIAQLAEEELKEDREILSLKQLEENITFS 120
Db 55 -----CASDQVEKCKVDIAQLAEEELKEDREILSLKQLEENITFS 95
Qy 121 KOIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALEROEYKELQKQELQSO 180
Db 96 KOIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALEROEYKELQKQELQSO 155
Qy 181 SILQOEKELSLARLQOQLCSFQEMTSERNVKEELKALAEALDAYOQKEEQSERLVKQLE 240
Db 156 SILQOEKELSLARLQOQLCSFQEMTSERNVKEELKALAEALDAYOQKEEQSERLVKQLE 215
Qy 241 EERKSTAEQLRLDNLRLREKEVELEKHIAHAQAAILIAOEKYNDTAQSLRDVTAQLESVQ 300
Db 216 EERKSTAEQLRLDNLRLREKEVELEKHIAHAQAAILIAOEKYNDTAQSLRDVTAQLESVQ 275
Qy 301 EKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLE 360
Db 276 EKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLE 335
Qy 361 SVQEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLE 420
Db 336 SVQEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLE 395
Qy 421 TAESTNQEYARMVDLQNRSTLKEEIKETSSFEKITDLDKNLQNRQDEDFRKQLEEK 480
Db 396 TAESTNQEYARMVDLQNRSTLKEEIKETSSFEKITDLDKNLQNRQDEDFRKQLEEK 455
Qy 481 KRTAEKENVMTLWELINKWRLLYEELYEKTPFOQLDAFAEAKOALLNEHGATQEQLN 540
Db 456 KRTAEKENVMTLWELINKWRLLYEELYEKTPFOQLDAFAEAKOALLNEHGATQEQLN 515
Qy 541 KIRDSYAQLLGHQNLKQIKHVVKLKDENSQKLSVSKLSRSLQVVRKQNELRLQELDKA 600
Db 516 KIRDSYAQLLGHQNLKQIKHVVKLKDENSQKLSVSKLSRSLQVVRKQNELRLQELDKA 575
Qy 601 LGIRHFDPSKAFCHASKENFTPLKEGPNCC 631
Db 576 LGIRHFDPSKAFCHASKENFTPLKEGPNCC 606
RESULT 6
AAU11438
ID AAU11438 standard; Protein; 713 AA.
XX AC AAU11438;
XX 12-MAR-2002 (first entry)
XX DE Rat hyaluronic acid binding protein RHAMM.
XX Rat; hyaluronic acid binding protein; RHAMM; gene therapy;
KW receptor for HA mediated motility; immunosuppressive; cytostatic.
KW conjugate; rheumatoid arthritis; scleroderma; liver fibrosis; cancer;
XX Rattus norvegicus.
OS

[illegible]

Db	404	EETKSTAEQLRRDDLLREKEIELEKNTAAHAQAQTVIAQEKYSDTAQTLRDVT-----	456
Qy	301	EKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLE	360
Db	457	-----	456
Qy	361	SVQEKYNDTAQSLRDVSAQLESYKSSSTLKEIEDLKLENLTLOEKVAMAEKSVEDVQOQIL	420
Db	457	-----	456
Qy	421	TAESTNOEYARMVQDQLQNRSTLKEEETKEITSSSFLEKITDLKNQLRQODEDFRKQLEEK	480
Db	500	TAESTNOEYAKVVQDQNSSTLKEAEITKETSSTLYBKITDLQNLRQODEDFRKQLEEK	559
Qy	481	KRTAEKENVMTLMEINKWRLLVEELYKTPFQOQLDAFAEAKQALLNEHCATQEQNL	540
Db	560	AKWTEKETATLMEINKWRLLVEELYKTPFQOQLDAFAEAKQALLNEHCATQEQNL	619
Qy	541	KIRDSYAQLLGHQNLKQIKHVVVKLKDENSQKSEVSKLRSQVLSQVLRKQKQNELRLQGLDKA	600
Db	620	KIRDSYAQLLGHQNLKQIKHVVVKLKDENSQKSEVSKLRSQVLSQVLRKQKQNELRLQGLDKA	679
Qy	601	LGIRHFDPSKAFCHASKENF---TPLKEGPNCC	631
Db	680	LGIRHFDPSKAFCHESKENVTTLKPLKEGPNCC	713
RESULT 7			
AAR43563			
ID	AAR43563 standard; Protein; 476 AA.		
XX	AAR43563;		
XX	05-APR-1994 (first entry)		
DT	Hyaluronan receptor.		
XX	Hyaluronan receptor; HA; RHAMM; mediated motility; wound; healing; diagnosis; treatment; cell locomotion; tumour invasion; birth defects; inflammatory disorder; Alzheimer's disease; dementia; Parkinson's diseases; Huntington's disease; AIDS; diabetes; auto; immune diseases; corneal dysplasia; hypertrophy; surgery; burns; strokes; multiple sclerosis; depression; schizophrenia; CNJ; contraception; in vitro fertilisation; embryo development.		
XX	WO9321312-A.		
PN	28-OCT-1993.		
PD	13-APR-1993; 93WO-CA00158.		
XX	09-APR-1992; 92GB-0007949.		
XX	(MANT-) MANITOBA CANCER TREATMENT & RES FOUND.		
PA	(UYMA-) UNIV MANITOBA.		
PI	Turley EA;		
PI	WPI; 1993-351722/44.		
DR	N-PSDB; AAQ51212.		
DR	DNA encoding hyaluronan receptor - used to produce proteins and antibodies for alteration of cell locomotion		
XX	Claim 7; Fig 23; 88pp; English.		
XX	The sequence is that encoded by a cDNA clone encoding the hyaluronan receptor (HARC). The sequence was obt'd. by screening a 373 library in lambda gtl with antibodies to HARC. A clone of 1.9 kb was obtained and used to rescreen the library to obtain the full length, 2.9 kb clone. HA is down regulated in stationary normal cells and is only expressed in situations where cell motility is desired, e.g. in wound healing, in response to growth factors and in chemotaxis by CC		

The sequence is that encoded by a cDNA clone encoding the hyaluronan receptor (HARC). The sequence was obtd. by screening a 373 library in lambda gtl with antibodies to HARC. A clone of 1.9 kb was obtained and used to rescreen the library to obtain the full length, 2.9 kb clone. HA is down regulated in stationary normal cells and is only expressed in situations where cell motility is desired, e.g. in wound healing, in response to growth factors and in chemotaxis by

FT		/note= "Region indicated in specification"
Binding-site	636..646	
FT		/note= "Hyaluronic acid binding domain"
FT	658..667	
Binding-site		/note= "Hyaluronic binding domain"
FT		
FT		
XX		
XX	WO9738098-A1.	
XX		
XX		
PD	16-OCT-1997.	
XX		
XX		

10-APR-1951, 3/MO-CA00240.
FF
XX

PK 10-APR-1956; 96GB-0007441.
XX
XX
PA (MANI-) MANITOBA CANCER TREATMENT & RES FOUND.
PA (UYMA-) UNIV MANITOBA.
PA

xx PI Entwistle J, Turley EA:

XX
DR .WPT: 1997-512715/47.

DR N-PSDB; AAV02800.
YY

Isolated human receptor for hyaluronic acid mediated motility - used to develop products for treating a tumours inflammatory

PT disorders, dementia, AIDS, diabetes and auto-immune diseases

PS Claim 16; Page 46; 66pp; English.

CC This sequence represents the human hyaluronan receptor which is also

CC Hyaluronan is a large glycosaminoglycan that is ubiquitous in the

CC migration, growth and transformation. It interacts with cell surfaces via

effects. The RHAMM/Hyaluronic acid interaction is involved in

used in the treatment of disorders involving cell locomotion, e.g., tumour

CC Alzheimer's and other forms of dementia, including Parkinson's and

CC functioning disorders, diabetes, autoimmune diseases, corneal dystrophies and hypertrophies, burns, surgical incisions and adhesions,

strokes and multiple sclerosis. They can also be used in e.g. CNS and spinal cord regeneration, contraception and in vitro fertilisation and

CC embryo development. The products can also be used in detection, diagnosis
CC and prognosis.

XX
50
Sequence 725 AA:

Query Match 62 98. Score 1053. DB 18. Length 725.

Best Local Similarity 64.8%; Pred: No. 2.2e-93;
Matches 411. Conservative 40. Mismatches 96.
Indels 99. Cons 3.

Matches 411; Conservative 49; Mismatches 88; Index 88; Gaps 3;

OV 1 MRALSLELMKLRNKRETKMRSMVMVKOEGMELKLOATOKDLTESKGIKIVOLEGKLSIEKE 60

db 164 LRILSLEIMKLRNKRETTKMRGMMAKOEGMEMKLOVTOBSLEESOGKIAOLEGKLVSTIEKE 223

OV 61 KIDEKCEKLLLEYIOEITSCASDOVEKCKVDIAOLEEDLKEKDREILSLKOSLEENIT-F 119

db 224 KIDEKSETKLEYYIEEISCASDOVEKYKLDIAOLEENLKEKNDEILSLKOSLEENIVIL 283

QY 120 SKQIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALERQEYEKLOQKELQS 179

db 284 SKOVEDLNVKCOLLEKEKEDHVNRNRHNNENI.NAEMONI.KOKFTI.EOOEHEKLOOKELOI 343

0Y 180 0SL00EKEISARI.000I.CSFOEEMTSEKNVKEFEIKIAEL.DAVOOKEE0SERI.VKOI. 239

344 DEI I OOEVEI CCEI HAVI CCEOEEMIKENI DEEI KOTI DEI DVI OOEVEAEPIKAI 103

Q: What was the date of the first meeting?

A: The first meeting was held on May 1st, 1968.

DB 104 FFFAVSDBFFI VI I FFFVI KCVFBAFI FVSSAAHTOATI I OFEYVDSMVIOSI FDNIM- --- 157

0-7 300 QZVYNDXACZY DRYEAOY ECEOEYVNDXACZY DRYEAOY 250

Db 458 ----- 457
QY 360 ESVQEKYNDTAQSLRDVSAQLESYKSTLKEITEDLKLNLTLQEKVAMAEKSVEDVQOOI 419
Db 458 -----AQFESYKALTASEIEDLKLENSLQEKAAKAGNAEDVQHOI 499
QY 420 LTAESTNOEYARMVDLQNRSTLKEEIEKITSSFLKTTDLKQLRQODEDFRQLEEK 479
Db 500 LATESNQEVYRMLDLQTKSALKETEIKETVSPFQKITDQLQNLKQOEEFRQLEDE 559
QY 480 GKRTAKENVMFELTMEINKWRLLYEELYEKTPFQOQDLDAFEAKQALLNHEGATQOOL 539
Db 560 EGRKAKEKNTTAEITMEINKWRLLYEELYNKTKPFQQLDQDAFEVEKQALLNHEGAQOOL 619
QY 540 NKIRDSYAOQLGHONLQKIKHVVKLDKENSOLKSVSKLRSOLVKRKONELRLOGELDK 599
Db 620 NKIRDSYAKLLGHONLQKIKHVVKLDKENSOLKSVSKLRSOLVKRKONELRLOGELDK 679
QY 600 ALGIRHFDPSKAFCHASKENF---TPLKEGNPNC 630
Db 680 VLGIKHFDPSKAFCHASKENFALKTPPLKEGNTNC 713

RESULT 9
ABG60842
ID ABG60842 standard; Protein; 725 AA.

XX AC ABG60842;
XX DT 13-AUG-2002 (first entry)
XX DE Human receptor for hyaluronan-mediated motility (RHAMM).
XX KW Tissue disorder; response-to-injury process; cell proliferating;
KW hyaluronic acid; HA; receptor for hyaluronan-mediated motility;
KW RHAMM; inflammatory neurological disorder; Parkinson's disease;
KW Alzheimer's disease; arthritis; multiple sclerosis; gastritis; nephritis;
KW inflammatory dermatosis; psoriasis; inflammatory bowel disease;
KW stenosis; restenosis; cancer; kidney fibrosis; inflammatory lung disease;
KW emphysema; asthma; cystic fibrosis; obesity; obesity related disease;
KW lupus; cardiovascular disease; atherosclerosis; wound; scar; diabetes;
KW tissue transplantation; stroke; inflammatory response; fibrotic response;
KW medical implant; Acquired immunodeficiency syndrome; AIDS; hepatitis;
KW myocardial fibrosis; hepatic fibrosis; chronic cystitis; acute mastitis;
KW septic shock; thyroiditis; retinopathy.

XX OS Homo sapiens.

XX PN WO200228415-A1.

XX PD 11-APR-2002.

XX PF 05-OCT-2000; 2000WO-IB01534.

XX PR 05-OCT-2000; 2000WO-IB01534.

XX PA (TRAN-) TRANSITION THERAPEUTICS & DIAGNOSTICS IN.

XX PI Turley EA, Cruz TF;

XX DR WPI; 2002-435298/46.

XX PT Treating tissue disorder associated with response-to-injury process or
PT proliferating cells in mammals, e.g. fibrosis, inflammation, by
PT administering a compound that alters activity of transition molecules
PT within a cell

XX PS Example 30; Fig 50; 215pp; English.

XX CC The invention describes a method of treating a tissue disorder associated
CC with response-to-injury process or proliferating cells in a patient,
CC comprising administering a polypeptide (I) which binds hyaluronic acid

CC (HA), an antibody which binds one of domains DI-D5 of Receptor for
CC hyaluronan-mediated motility (RHAMM), a polypeptide fragment encoding
CC any of DI-D5 of RHAMM, or a vector which expresses antisense RHAMM,
CC antibodies or a polypeptide fragment. The method is useful for treating a
CC patient with an inflammatory neurological disorder such as Parkinson's
CC disease, Alzheimer's disease, arthritis including rheumatoid arthritis,
CC osteoarthritis, multiple sclerosis, inflammatory dermatosis (psoriasis),
CC inflammatory bowel disease, stenosis or restenosis, cancer, kidney
CC fibrosis, inflammatory lung disease (e.g. emphysema, asthma, cystic
CC fibrosis), obesity or obesity related diseases, lupus, cardiovascular
CC disease (e.g. atherosclerosis), and wound especially surgical excision
CC adhesions, to prevent scar and also for treating or preventing diabetes
CC mellitus. The method is also useful for treating tissue transplantation
CC (e.g. skin grafts), stroke, inflammatory responses or fibrotic response
CC associated with medical implants such as hip implants, vascular wraps and
CC catheters), inflammatory diseases such as AIDS, myocardial and hepatic
CC fibrosis, chronic cystitis, acute mastitis, gastritis, nephritis,
CC hepatitis, septic shock, thyroiditis, and retinopathy. This sequence
CC represents a receptor for hyaluronan-mediated motility protein used in
CC the method of treating a tissue disorder described in the invention.

XX XX Sequence 725 AA;

Query Match 62.9%; Score 1953; DB 23; Length 725;
Best Local Similarity 64.8%; Pred. No. 2.2e-93;
Matches 411; Conservative 49; Mismatches 86; Indels 88; Gaps 3;

QY 1 MRALSLELMKLRNKRRTKMRSMVMVKQEGMELQLOATOKDLTSKGVQLECKLVSIKIE 60
Db 164 LRLSLELMKLRNKRRTKMRSMVMVKQEGMELQLOATOKDLTSKGVQLECKLVSIKIE 223
QY 61 KIDECETEKLEYLEIQEISCASDQVEKCKVDIAQLEEDLKEKDREITLSLQSLNIT-F 119
Db 224 KIDECETEKLEYLEIEEISCASDQVEKCKVDIAQLEENLKEKNDITLSLQSLNITVL 283
QY 120 SKQIEDLVKCOLLETERDNLVSKDRERATLSAEMQILTERLALERQBYEKLOQELQS 179
Db 284 SKQVEDLVKCOLLEKEKEDHVNRRNHNENLNAEMONLQKFIQEQSHKLOQELQI 343
QY 180 QSLQOQEFSLARLOQLCSFOEEMTSEKNVFEELKALAEADAVQOKEQSERLVKOL 239
Db 344 DLSLQOQEFSLARLOQLCSFOEEMTSEKNVFEELKALAEADAVQOKEQSERLVKOL 403
QY 240 EEEKSTAEQTLRLNLLREKEVELEKHAHAQAAILTAQEKYNDTAQSLRDVTAQLESV 299
Db 404 EEEKSTAEQTLRLNLLREKEVELEKHAHAQAAILTAQEKYNDTAQSLRDVTAQLESV 457
QY 300 QEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAOL 359
Db 458 ----- 457
QY 360 ESVQEKYNDTAQSLRDVSAQLESYKSTLKEITEDLKLNLTLQEKVAMAEKSVEDVQOOI 419
Db 458 -----AQFESYKALTASEIEDLKLENSLQEKAAKAGNAEDVQHOI 499
QY 420 LTAESTNOEYARMVDLQNRSTLKEEIEKITSSFLKTTDLKQLRQODEDFRQLEEK 479
Db 500 LATESNQEVYRMLDLQTKSALKETEIKETVSPFQKITDQLQNLKQOEEFRQLEDE 559
QY 480 GKRTAKENVMFELTMEINKWRLLYEELYEKTPFQOQDLDAFEAKQALLNHEGATQOOL 539
Db 560 EGRKAKEKNTTAEITMEINKWRLLYEELYNKTKPFQQLDQDAFEVEKQALLNHEGAQOOL 619
QY 540 NKIRDSYAOQLGHONLQKIKHVVKLDKENSOLKSVSKLRSOLVKRKONELRLOGELDK 599
Db 620 NKIRDSYAKLLGHONLQKIKHVVKLDKENSOLKSVSKLRSOLVKRKONELRLOGELDK 679
QY 600 ALGIRHFDPSKAFCHASKENF---TPLKEGNPNC 630
Db 680 VLGIKHFDPSKAFCHASKENFALKTPPLKEGNTNC 713
RESULT 10

AAU11436	
ID	AAU11436 standard; Protein: 725 AA.
XX	
AC	AAU11436;
XX	
DT	12-MAR-2002 (first entry)
XX	
DE	Human hyaluronic acid binding protein RHAMM.
XX	
KW	Human; hyaluronic acid binding protein; RHAMM; gene therapy;
KW	receptor for HA mediated mobility; immunosuppressive; cytostatic.
KW	conjugate; rheumatoid arthritis; scleroderma; liver fibrosis; cancer;
XX	
OS	Homo sapiens.
XX	
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 278 /note= "Encoded by GNA"
FT	Misc-difference 299 /note= "Encoded by AAA"
FT	Misc-difference 323 /note= "Encoded by AAA"
FT	Misc-difference 331 /note= "Encoded by CAG"
XX	
PN	WO200180899-A2.
XX	
PD	01-NOV-2001.
XX	
XX	
PF	20-APR-2001; 2001WO-CA00533.
XX	
PR	20-APR-2000; 2000US-198613P.
XX	
PA	(CANG-) CANGENE CORP.
XX	
PI	Woloski BMR, Williams AM, Sereda TJ, Wiebe DJ;
XX	
DR	WPI; 2002-075094/10.
DR	N-PSDB; AAS17496.
XX	
PT	Protein conjugates that selectively target certain tissues and organs
PT	useful for treating and preventing various diseases, comprises
PT	glucose-aminoglycan-targeting domain conjugated to a therapeutic
PT	protein -
XX	
PS	Claim 6; Page 115; 121pp; English.
XX	
CC	The invention relates to a conjugate comprising an hyaluronic acid (HA)
CC	-binding protein e.g. RHAMM (receptor of HA mediated mobility) or peptide
CC	contiguous with, or coupled to a polypeptide conjugated to a therapeutic
CC	agent, and the polynucleotides encoding them. Also included is a method
CC	for preparation of the HA-binding protein by inserting a first nucleotide
CC	sequence encoding a HA-binding protein directly linked to a second
CC	nucleotide sequence encoding a therapeutic protein into a suitable
CC	vector, expressing the vector in an acceptable host, purifying conjugate
CC	molecule from host or expression medium. The composition is useful for
CC	altering in vivo the distribution of a therapeutic agent comprising
CC	administering the composition to the animal where conjugate molecule will
CC	distribute primarily in tissues and organs containing high levels of
CC	endogenous HA and for treating mammal with a disorder where a diseased
CC	tissue of the mammal contains high level of HA e.g. rheumatoid
CC	arthritis, scleroderma, liver fibrosis and cancer. Lower therapeutic
CC	dosages required also translates into lower immunogenicity of the
CC	conjugated protein as compared to the native protein. As a result,
CC	conjugates improve patient compliance and reduce direct and indirect
CC	costs associated with the drug substance and its administration.
CC	Conjugates allows for the use, where appropriate, of lower, safer,
CC	dosages as compared to the conventional dosage requirements for the
CC	unconjugated corresponding therapeutic agent. Conjugate molecules have an
CC	increased half-life and potency, resulting in prolonged circulation of
CC	the molecule, efficient distribution into the target tissues, and
CC	increased bioavailability. The present sequence represents a RHAMM
CC	protein.

Db 1013 TLDDQLQMEDQVNTLTAKTKLEQVDDLEGSLEQEKKLCMDLERAKRKLGLDKLAQES 1072
 Qy 221 ELDAVQOQEEQSERLVKQ-----LE---EERKSTAEOLTRLDNLREKEVELEKHAH 271
 Db 1073 TMDTENDQOQNEKLKKEFEMSNQGLKEDQALAMQOLKIKELQARIIELEEEIEAE 1132
 Qy 272 AQAILIAQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESQEKYNDTAQSL 331
 Db 1133 RASRAKAEKQSDLSRELEIEISERLE-----EAGGATSAQIELNKKREAEFOKMR 1182
 Qy 332 RDVTAQLESQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESQEKYNDTAQSL 391
 Db 1183 RD-----LEESTLQHEATAAALRKHHA--DSVAE-IGKQIDSLQRYKQKLEKEKSELKMEI 1235
 Qy 392 EDLKLNTLTQEKVAMAEKSVEDVQOQILTAESTNQEVARMVQDLONRSTLKEEIEKIT 451
 Db 1236 NDLASMETVSKAKANPEKMCETLEDQLESEIKTKEEEOORLINELSAQARLHTEGFEFS 1295
 Qy 452 -----SSFLEKIDTLKNLRQO-----DED-FRKQLE 477
 Db 1296 RQLEKDAWVSQSRGKQFTQOIEELKQLEETKAKSTLAHALQASARHDCDLLREQYE 1355
 Qy 478 EKGKTAENVMETLTWEINKWRLY-----EELYEKTTPQOQLD-----519
 Db 1356 EQEAKAELQRMGSKANSEVAQWRTKYETDAIQRTEELEEAQKLAQRLQDAEHEVAVN 1415
 Qy 520 -----AFEAQKQALLNE-----HGATQEOQLN 540
 Db 1416 SKCASLETKQRLQNEVEDLMIDVERSNACIALDKKORNFQKVLAEWKQRYEETQAELE 1475
 Qy 541 -----KIRDSYAQLLGH-----ONLKOKI-----KHVYKLD 567
 Db 1476 ASQKESRSLSTELFKVKNAYEESLDHLETLKRENKLAQOEISDLTEQIAEGGKHTELEK 1535
 Qy 568 ENSQLKSEVSKLRSOLVRKQNELRLOGLDYLKALGIR 604
 Db 1536 VKQLDHEKSELOTSL-----EEAASLEHEEGKILRIQ 1569

RESULT 13

AAW02258
 ID AAW02258 standard; Protein; 1411 AA.
 AC AAW02258;
 XX
 DT 09-MAR-1997 (first entry)
 XX
 DE Nucleolar/endosomal auto-antigen p162.
 XX
 KW Auto-antibody; p162; rheumatic disease; antigen; diagnosis;
 KW gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN DE19515514-C1.
 XX
 PD 12-SEP-1996.
 XX
 PF 27-APR-1995; 95DE-1015514.
 XX
 PR 27-APR-1995; 95DE-1015514.
 XX
 PA (PRIV-) PRIVATES INST IMMUNOLOGIE & MOLEKULARGEN.
 XX
 PI Renz M, Seelig HP;
 XX
 DR WPI; 1996-403153/41.
 DR N-PSDB; AAT58751.
 XX
 PT DNA encoding nucleolar-endosomal auto-antigen - useful for exact
 PT diagnosis of rheumatic disease, in gene therapy and for removal of
 PT specific auto-antibodies
 XX

PS Claim 1; Fig 2; 15pp; German.
 XX
 CC Transformed cells can be cultured to produce the antigen p162, for use
 CC in exact (differential) diagnosis of rheumatic disease, i.e. they
 CC can detect, in immunoassays, Western blots, etc., rheumatic-
 CC specific auto-antibodies. The antigen can be used therapeutically,
 CC in the removal of auto-antibodies from the circulation, or when
 CC coupled to a cytotoxin, the elimination of auto-antibody-
 CC producing lymphocytes.
 XX
 SQ Sequence 1411 AA;
 Query Match 13.3%; Score 414; DB 17; Length 1411;
 Best Local Similarity 23.0%; Pred. No. 1.5e-13;
 Matches 168; Conservative 150; Mismatches 220; Indels 192; Gaps 28;
 Qy 7 ELMKLRN-----KRETKMRSMVMYKQEG--MEIK-----LQAT--OKDLTSEK 44
 Db 289 ELQKLKSSVNEITQKNQTLTENLKKKEQDYTKLEEKHNEESVSKKNIQATLHKQDL---- 344
 Qy 45 GKIVOLEGLVSIKE-----KIDKCE-TEKLLYEIQEISCASDOVEKCKVDIAQLEE 97
 Db 345 -DCQQLQSRRLSASETSLRHIVHVELSEKGEATQKLEELSEV---ETKQHLKAEFKQLQ 400
 Qy 98 DLKEKDREILSKQSLSEENITFSKOIEDITVKCOLLETEDNLVSKDRERAEFTLSAEMQI 157
 Db 401 QREKEQHGLQLOSEIQ-----LHKKLETERQLGEAHR-----LKEQRL 443
 Qy 158 LTERALAREQYEKLOQKELQSLQLOQKELSAKLQOOL-----CSFQEMTSE 207
 Db 444 SSEKLMDEQVQVADQLKLSLEQLKEKVTNSTELQHLQDKTKQHQHQQALQOSTTAK 503
 Qy 208 KNVFEELKLALAEADVAQOKEEQSERLVQ-----LEERK-----STA 247
 Db 504 LREAQNDLEQVLRQIGDKDKQKIONLEALQKSKENISLEKEREEDLYAKIQAGEETAVL 563
 Qy 248 EQLTRLDNLLREKEVELEKHAHQAAILIAQEKYNDTAQSLRDVTAQLESVQEKYNDTA 307
 Db 564 NQLQEKNHTLQEQVTQTLTEKLNQSESHKQAOENLHDQVOEQK---AHLRAAQDRVLSLE 620
 Qy 308 QSLRDVTAQLESQEKYND-----TAQSLRDVTAQLESQEKYNDT 348
 Db 621 TSVNELNSQLNESKEKVSQLDIQIKAKTELLLSAFAAKTAQR-ADLQNLHDLTAQNALQDK 679
 Qy 349 AQSLRDVTAQLESVQEKYNDTAQSLRDVSAQLESYKSSTL-----KLEIDL-----KLENT 400
 Db 680 QOELNKIITQDQVTAKLQDKQEHCSQLESHLKYEKYLSEQKTEEGQIKKLEADS 739
 Qy 401 LOEKVAMAEKSVEDVQOQILTAESTNQEVARMVQDLONRSTLKEEIE-----KEITSF---- 454
 Db 740 LEVK-ASKEQALQDLQOQ-----RQLNTDLELRATELSKQLEMEKEIVSSTRLD 787
 Qy 455 ----LEKITDNLKNLRQODEDFRKQLEKGTAKENVMETLTWEINKWRLLYELYEK 510
 Db 788 LQKSEALESIKQKLTQKEE--KQILKQDFETLSQET-----KIQHEELNMR 833
 Qy 511 TKPFOQLDAFAEAKQALLNEHGATQEOQLNKIRDSY-----AQLLGHONL 555
 Db 834 IQTTVTELQKVAREALMTLSTVTKDLKSRVSDSKNSKSEFEKENQKGAAILDLEKT 893
 Qy 556 KQIKHVVKLKDENS-----OLKSEVSKLRSOLVKR-----KON---- 589
 Db 894 CKELKHQLOVQMENTLKQKELKKSLEKEKASHQKLKLELNSMQLOLQAOQNTLQKNEKE 953
 Qy 590 ELRLQGLDK 599
 Db 954 EOQLQGNINE 963
 RESULT 14
 ABB61144
 ID ABB61144 standard; Protein; 1690 AA.
 XX

THIS PAGE BLANK (USPTO)